


```
QY 143 TCACAACTGTCTCTCGACACATCTCTG-----CAATCCGGCACACC 187
Db 52 LysHisGlySerSerAspGlyLeuIleSerAlaIleAsnAlaGlnSerHisHis 71
QY 188 CTTCTATAACAACGGGTATCCGACATCTACCTGCTCGACGACCCCCACGACCCCTTGGC 247
Db 72 ProAsnGlnSerGly-----TyrTyrProIleThrThrGlyAlaAsnAlaPheAla 88
QY 248 GCCCGCGCCGCTTATCGAATCTCGCGAACACAGCTCGATTTGCAATPACTACATTTGG 307
Db 89 AlaArgSerThrLeuThrAspValAlaAsnGlnSerIleAspIleGlnTyrIleTrp 108
QY 308 CCAACGACATTTCCGGAGGTGCTGTTCAACCTCATGTACTTCCGCGACAAAGCGCGC 367
Db 109 HisAsnAspGluAlaGlyGlnLeuMetLeuLysAspLeuTrpGluAlaAlaAspArgGly 128
QY 368 GTGCGGTACGCTGCTGTGTGACGAC---AACACACGCGGGTGTGACCATCTCTCTG 424
Db 129 ValIleValArgLeuLeuLeuAspPheAsnSerSerProGluLeuAspGlnLeuLeu 148
QY 425 CTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCAACCCCTCTCTCTACGC 484
Db 149 LeuArgIleSerLysHisLysAsnIleAlaValArgLeuIleAsnProMetProTyrArg 168
QY 485 AAATGGCGGCATCTGGCTACTCTGACGACTTCCCGCCTCAACGGCGCGATGCACAA 544
Db 169 GlyPheArgSerLeuAsnTyrMetLeuHisProIleArgThrAsnArgArgMetHisAsn 188
QY 545 AAATCTCTTACCGCGACAAACCGCGCACCATCTCGCGGCGAGCAATATCGCGACGAA 604
Db 189 LysSerMetThrPheAspAsnLysIleSerValIleGlyArgAsnIleGlyAsnGlu 208
QY 605 TACTTCAAAGTCGTGAGACACCGTTTTCGCGACTCGACATCTCTGCCACCGCGACG 664
Db 209 TyrLeuAsnAlaProValAsnHisPheAlaAspLeuAspValMetLeuValGlyHis 228
QY 665 GTGCTGGGGAAGTATCGACACTTCGACCGCTACTGGGAGCGCATTCGCCCCACAA 724
Db 229 ValValGlyIleThrGlnSerPheGluIleTyrTrpAlaSerProLeuSerPheAsp 248
QY 725 GCCACGCGCATCATCGCAGCGCAACATCGGC-----AAGGTCTTCAAGCATCGGA 778
Db 249 IleGluThrLeuValLysHisAspAsnLysAspAspIleSerGlyIleLysProValVal 268
QY 779 TACAACGAC-----GAAACATCCAGACACGCGCTCTCTGCGCTAC 817
Db 269 PheAspGluLeuLysValLysAspSerSerAspAlaAspArgGluLeuArgThrTyr 288
QY 818 CGCGAAACGTCGAACAGTCCGCCCTCTACCAAAATAACAGCGGACGATCGACTGG 877
Db 289 ArgGlnAlaMetGlnAsnSerThrIleGlyGlnAspLeuLeuAlaGlnValProPhe 308
QY 878 CAGAGCGTCCAAACCGCTGATCAGCGACACCCCTGCAAAAGAGCATCGACCGCGACC 937
Db 309 PheTrpThrGluIleAspLeuIleAlaAspAsnValIleLysLeuAsnGlyHisSerAsp 328
QY 938 CGCAAAACCGCGATTCGCGGAGGTGGAAGCGCCTCAACAGCCGCAAAAGCGTC 997
Db 329 ProAsnGluPheLeuValSerGlnLeuGlnThrGlnLeuGlyGlnProSerLysLysLeu 348
QY 998 TATCTGTTTCACTTCTGTCCTTACAAATCCGGCACAGCGACTGCGCAAAACTG 1057
Db 349 SerIleIleSerSerTyrPheValProThrLysGluGlyIleAspThrLeuIleThrLeu 368
QY 1058 GTGACGAGCGCATAGACGTTTACCGTCTGACCACTCGCTACAGCGCGACCGAGTTGCC 1117
Db 369 AlaLysMetGlyValGluValArgIleLeuThrAsnSerPheAspAlaThrAspValGly 388
QY 1118 GCGTCCATTCGGCTACGTCAAAATACGAAACCGCTGCTCAAGCGCGCATCAAACTC 1177
Db 389 IleValHisAlaGlyTyrAlaHisTrpArgLysGlnLeuLeuAlaAlaGlyValHisLeu 408
QY 1178 TAGACACTGCAACCCCAAC---CATGCGGTCCCGCGCACAAAGACAAA-----GCCTG 1228
```

```
Db 409 PheGluIleLysSerSerAlaGlnSerIleGlnAspAsnGluAsnArgPheTrpArgThr 428
QY 1229 ACCGCGACTCCGTAACACGAGCTGCATGCCAAAACCTTTTCATTGTGGACGGCAACGCAATC 1288
Db 429 ArgGlnHisSerThrThrSerLeuHisAlaLysAlaPheAlaValAspAspGlnIle 448
QY 1289 TTCATCGGCTCATTCACCTCGACCCCGTTCCGCGACGCTCAATACCGAAATGGCGTC 1348
Db 449 PheIleGlySerTyrAsnValAspProArgSerAlaAsnIleAsnThrGluLeuGlyVal 468
QY 1349 GTCATCGAAAGCCCAAAATCGAGAACAGATGGAGCGCACCTCGCGATACC----- 1402
Db 469 LeuIleLysAspSerLysLeuAlaGlyGlnLeuHisLysAlaLeuSerAsnSerGlnAla 488
QY 1403 ACACCGCAATACGCTACCGGTTTACCTCGACAAACACAAACCGCTGCAATGGCAGCAT 1462
Db 489 IleThrHisGlnAlaTyrGluLeuLysLeuAspAlaLysGlyAsnIleAsnTrpHis--- 507
QY 1463 CCGCGCACCCGAAACAACTACCGGAC-----GAACCCGAA 1498
Db 508 -----ThrIleGluAsnGlyGlnProValIleLeuHisHisGluProHis 522
QY 1499 GCCAAACTTTGGAAACGATCGCGCAAAAATCCTATCCCTGCTGCCCATCGAAGGTTTA 1558
Db 523 MetAsnHisLysAspArgIleIleIleTrpLeuAlaGlyLeuMetProIleAspTrpLeu 542
QY 1559 TTA 1561
Db 543 Leu 543

RESULT 2
US-09-252-991A-23909
; Sequence 23909, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23909
; LENGTH: 564
; TYPE: PRX
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23909

Alignment Scores:
Pred. No.: 1e-54 Length: 564
Score: 790.00 Matches: 200
Percent Similarity: 52.5% Conservative: 86
Best Local Similarity: 36.7% Mismatches: 189
Query Match: 27.7% Indels: 70
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x US-09-252-991A-23909 (1-564)
QY 47 CCGACGCTCATTTCCCTTTTATGCTCTCTCTCTCATGTCTTCATGTTGCCCA 106
Db 50 ArgArgPheLeuLeuLeuAlaLeuLeuAlaLeuSerGlyCysAlaSerThrProPro 69
QY 107 -----CTGGAAGAAGCGACGGAAGCCGT 130
Db 70 GlnProSerSerAlaLeuProAlaGluGlyThrTrpLeuAlaArgGlnAlaGlu----- 87
QY 131 CATTTCAATACTTCCAAACCTGCTCTCTGGCAACATCTGCAAAATCCGGCACACCCCT 190
Db 131 -----
```


Db 88 -----IleGlnGlyArgAspHisPro 94
QY 191 CATAACAGCGGTATCCGACATCTACTGCTGCGAGCACCACCGAGCCCTTGGCGCC 250
Db 95 -----GlyGlnSerGlyPheHisLeuLeuGlyAlaSerGluAspAlaPheValAla 111
QY 251 CGCGCGCCCTTATCGAATTCGCGAACAACACGCTCGATTGCAATCTACTATTTGGCGC 310
Db 112 ArgAlaAlaLeuIleArgAlaAlaGlnArgSerLeuAspIleGlnTyrTyrIleValHis 131
QY 311 AACGACATTTCCGCGAGCTCTGTTCAACCTCATGTCTGCGGAGAACGCGGGTG 370
Db 132 AspGlyLeuThrArgAlaLeuAlaTyrGluLeuLeuIleAlaAlaAspArgGlyVal 151
QY 371 CGGTAGCCCTGCTGTTGGACGACAAACACGCGCGGTGGACGATCTCTGCTGCC 430
Db 152 ArgValArgIleLeuIleAspAspThrAlaSerAspGlyTyrPheTyrGluIleGlyVal 171
QY 431 CTCGACAGCCATCCCAATATCGAAGTGCCTGTTCACCCCTTCTGCTACGCAAA--- 487
Db 172 LeuSerAlaHisProAsnIleGlnValArgLeuPheAsnProLeuHisLeuGlyArgAla 191
QY 488 -----TGGCGGCACTCGGCTACTGACCGCACTTCCCGCCCTCAACCGCGCATG 538
Db 192 ThrGlyLeuThrArgGlyValGlyArgLeuPheAsnLeuSerGlnHisArgArgMet 211
QY 539 CACAACAAATCCTTTACCGCGCAACACCGCGCCACCATCTACGCGAGCAATATCGC 598
Db 212 HisAsnLysLeuTyrPheAlaAspGlyThrAlaAlaIleValGlyGlyArgAsnLeuGly 231
QY 599 GACGAATCTTCAAGTCGGTGAGGACACCGTTTTCGCGCACTGACATCTCGCCACC 658
Db 232 AspGluTyrPheAsnAlaLysProGluMetAsnPheThrAspLeuAspLeuLeuGlyVal 251
QY 659 GGCAGCGCTCGCGGGAAGTATCGCAGCACTTCCACCGCTACTGGGCAAGCATTCGCC 718
Db 252 GlyProIleAlaAsnGlnLeuSerHisPheAspGlnTyrTrpAsnSerAlaIleSer 271
QY 719 CACAACGCCACG-----CGCATCATCGCGAGCGCAACATCGCGAGGCTT 766
Db 272 ArgProIleGluAspPheLeuTyrPheArgAlaProTyrProGlyGluLeu----- 287
QY 767 CAAGCATCTGGATACAACGAGAAACATCAGACCGCGCTCTGCGCTAC-----CGC 820
Db 288 -----GluSerAlaArgLysLeuGlnArgTyrLeuArgLys 300
QY 821 GAAACCGTCGAACAGTCG-----CCCTCTACCAAAATAACAGACGGGACGC 868
Db 301 GluSerValLysGluSerGlyTyrIleArgHisLeuPheAspArgGlyAspGlnProArg 320
QY 869 ATC-----GACTGGCAGAGC-----GTCCAAACCGCGCTGATCAGCGACACCCCT 913
Db 321 LeuGlyAsnTyrPheGluAsnLeuThrTrpAlaArgAlaGluAlaIleTyrPheAlaPro 340
QY 914 GCAAAAGGATCGACCGCGACCGCCGCAAAACCGCG-----ATTGCGGGAGCGCTCAA 967
Db 341 LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuSerProHisLeuAla 360
QY 968 GACGCGCTCAACAGCCCGAAAAGCGTCTATCTGTTTCACCTATTTCGTCCTTACA 1027
Db 361 GlyLeuPheLysGlyValGlnLysGluLeuIleLeuValSerAlaTyrPheValProAla 380
QY 1028 AAATCCGGCACAGACGCACTGGCAAAACTGGTGAGGACGGCATAGACGTTACCGCTCG 1087
Db 381 LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu 400
QY 1088 ACCAATCTGTCAGCGGACCGAGTTCGCGCGCTCCATTCGCGCTACGTCGAATACCGA 1147
Db 401 ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg 420
QY 1148 AAACCGCTGCTAAAGCCGCACTCAAACTCTAGAGCTGCAACCAACATCGCGTCCCC 1207
Db 421 MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro 440

QY 1208 GCCACA---AAGACAAAGCGCTGACCGGAGCTCCGTAACCGAGCTGCTGCAAAACC 1264
Db 441 LeuSerGlyAlaProTyrArgLeuHisGlySerSerSerAlaSerLeuHisSerLysAla 460
QY 1265 TTCATTGTGAGCGCAAAACGATCTTCATCGGCTCATTTCAACCTCGACCCCGTTCCGCA 1324
Db 461 MetValPheAspArgArgLysValPheIleGlySerPheAsnPheAspProArgSerIle 480
QY 1325 CGGCTCAATACCGAATGGCGTCTCATCGAAGCCCAAAATCGGAGAACAGATGGAG 1384
Db 481 LeuTrpAsnThrGluValGlyValIleValAspSerProLeuLeuAlaGluGlnValArg 500
QY 1385 CGC---ACCCTCGCGGATACACACCGAATAGCCCTACCGCTTCCCTCGCAAAACAC 1441
Db 501 GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer 520
QY 1442 AAC-----CGCTGCAATGGCAGAT-----CCCGCCACCCGAAACCTACCCG 1486
Db 521 GlySerArgProLysLeuValTyrIleAspGluArgAspGlyArgProGlnValLeuArg 540
QY 1487 AACGAACCCGAAACCACTTTGGAACGATCGCGCAAAATCTATCCCTGCTGCC 1546
Db 541 HisGluPro---GlySerLeuTyrArgLeuAsnAlaTyrValAlaGlyMetIleGly 559
QY 1547 ATCGAAGGTTTATTA 1561
Db 560 LeuGluLysMetLeu 564

RESULT 3
US-09-328-352-5559
; Sequence 5559, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5559
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5559

Alignment Scores:
Pred. No.: 3,59e-31 Length: 271
Score: 493.50 Matches: 102
Percent Similarity: 63.9% Conservative: 47
Best Local Similarity: 43.8% Mismatches: 73
Query Match: 17.3% Indels: 11
DB: 2 Gaps: 4

US-10-665-990A-13 (1-1561) x US-09-328-352-5559 (1-271)

QY 59 TCCCTTTATGCTCTCTCTGTTTCATGTTCTTATGTTGCTCC----- 103
Db 37 SerSerPheAlaValAlaLeuThrGlyCysSerThr---LeuProLysHisSerProGlu 55
QY 104 CCATCTGGAAGACGGAGCGGAAAGCGCTCATTTCAATACTTCCAAACCTGCTCTCTGGAC 163
Db 56 ProIleGln-----TyrAlaArgAspIleAspThrSerGlnThrSerLeuSerLys 72
QY 164 AACATCTCTCAATCCGGCACACCCCTCATACACACGGGCTATCCGACATCTACCTGCTC 223
Db 73 IleIleThrProLeuArgGlu-----LysAsnProAsnLeuThrGlyTyrHisLeuLeu 90
QY 224 GAGGACCCCAAGACCCCTTGGCGCGCGCGCTTATCGAATCTCGCGAACACAGC 283
Db 91 AsnAspProLeuGluAlaLeuAlaIaArgLeuArgLeuIleAspLysAlaGluLysThr 110

Db 62 ----- 62
QY 188 CCTCATAACAAGGGCTATCCGACATCTACCTGCTCGACGACCCCGCCAGAGCCCTTGGC 247
Db 63 -----GlnGlyLeuThralapheValalaleuAspAlapheMetSerileAla 79
QY 248 GCCCGCGCGCTTATCGAATCTCGGAACACACAGCTCGATTTGGCAATACATATTGG 307
Db 80 SerArgileTyLeuileArgAsnAlaLysGluThrlleAspLeuGlnTyTyrlleTyr 99
QY 308 CGCAACGACATTTCCGCGAGCTGCTGTTCACCTCATGTACCTTGGCGAGAGCGGCG 367
Db 100 ThrAsnAspPheValGlyAsnLeuileuHisGluLeuLeuLysAlaAlaAspArgGly 119
QY 368 GTGCGCGTACCTGCTGCTTCCGACAGCAACACGCGCGGTGGAGCATCTCCTGCTC 427
Db 120 IleLysValArgLeuLeuileAspAspGlnhnglylleLysLeuAspGlylleuArg 139
QY 428 GCCCTCGACAGCCATCCCAATATCGAAGTGCCTGCTTCAACCCCTTCTGCTACGCAAA 487
Db 140 SerLeuLeuGlnHisThrAsnPheluleArgLeuPheAsnProTyTyLysPheArgTyr 159
QY 488 TGGCGCGCATCTCGCTACCTACGAGCTTCCCGCGCTCAACCGCGCATGACACAA 547
Db 160 LeuArgilePheAspTyLeuPheArgPheLysLysValAsnHisArgMetHisAsnLys 179
QY 548 TCCTTTACCGCGGACAAACCGCGCCACCATCTCGCGGAGCGCAATATCGCGACGATAC 607
Db 180 LeulleleAlaAspAlaSerileAlavalThrGlyArgAsnIleSerSerGluTyr 199
QY 608 TTCAAAGTCGTGAGGACACCGTTTTCGCGACCTGACATCTCGCGCACCGGAGCGTC 667
Db 200 PheGluAlaSerSerLysPheGlnPheThrAspMetAspileLeuPheTyGlyHisAla 219
QY 668 GTCGCGCAAGTATCGCAGCATCTCGACGCTACTGGGCAAGCCATTCGCGCCACAGCC 727
Db 220 ValArgHisAlaGlnAlavalPheThrAspPheTyGluSerThrLeuSerValAsnAla 239
QY 728 ACGCGCATCTCCGAGCGCAACATCGGCAAG--GGTCTTCAAGCACTCGGATACAAC 784
Db 240 ThrGluillele-----GlyThrCysAlaGluHisLeuLysAlaLeuArgGluHis 257
QY 785 GACGAACA-----TCCAGACGCGCTCTGCGCTACCGCGAAACCGTCGACACAG 835
Db 258 TyrGluGlnLeuHisGluAspHisSerLeuThrGluAspLysLeuTyAspAlaGln 277
QY 836 TCGCCCTCTACCAAAAAATACAGCGGAGCGATCGACTGGCAGAGCGTCCCAACCCGC 895
Db 278 SerTyLeuLysGluLeuLeuGluHisAsnProilleGlnTrpSer-----LysAlaHis 295
QY 896 CTGATCAGCGACACCCCTGCAAAA 919
Db 296 PheValalaAspSerProLysLys 303

RESULT 6
US-09-538-092-1258
; Sequence 1258, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIORITY FILING DATE: 2000-03-29
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapadSeqFormatter Version 0.9
; SEQ ID NO 1258
; LENGTH: 5179

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q02817
US-09-538-092-1258
Alignment Scores:
Pred. No.: 3,86e-18 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 2 Gaps: 22
US-10-665-990A-13 (1-1561) x US-09-538-092-1258 (1-5179)
QY 119 ACGGAAAGCGCTCATTTCAAT-----ACTTCCAAACCTGCTCTCTCGACAAC 166
Db 1248 ThrValglulysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTCTCAAAATCCGGCACACCCCTCATPAACAACGGGCTATCCGACATCTACCTCGCAC 226
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
QY 227 GACCCCGACAGGCCCTTGGCGCGCGCGCTTATCGAATCTGCGAACACAGCGCTC 286
Db 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTCGAATACATACATTTGGCGCAACGACATTTTCGGCAGGCTGTGTTCAACCTCA-- 344
Db 1293 rValLeuSerThrTh-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACCTTCCCGCAGAACCGCGGCTGCGGTACGCC----- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheileCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyAspTyTyLysileArgValAs 1387
QY 381 -TCCTGTTGG-----ACGACAACAAC 400
Db 1387 nCysCysTrpProMetAspLysCysileThrThrProSerProProThrThrThrProSe 1407
QY 401 ACGCGCGGTTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCATGAAGTCGC 460
Db 1407 rProProProThrThrThrThrLeuProProThrThrThrPro----- 1422
QY 461 CTGTTCAACCCCTTCTGCTACGCAAAATGGCGCGCATCGGCTACCTGACCGACTTCCCC 520
Db 1423 ---SerProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1441
QY 521 CGCTCAACCGCGCATGACACAATAATCTTTACCGCGCAGACACCGCGCCACCATCTC 580
Db 1441 oileThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGCGAGCAATATCGCGACGAATACTTCAAAGTCGGTGAGGACACCGTTTTCGCGCAC 640
Db 1457 -----ProIleSerThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1474
QY 641 CTGACATCTCGCCACCGCGAGCGTGTGCGGCAAGATATCGCACGACTTCGACCC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrThrPr 1494

Db 644 SerThrThrGluSerSerSerAlaProValProThrProSerSerSerThr-----Thr 661
 Qy 1091 AACTCGTGTACAGCGACCGACGAGTGTGCGCG-----TCCAT 1126
 Db 662 GluSerSerAla-ProValProThrProSerSerSerThrThrGluSerSerSerAla 681
 Qy 1127 TCCGGCTACGTCAAAATACCGAAAACCGCTGCTCAAAGCCGGCATCAAACTCTACG----- 1181
 Db 681 aProValThrSerSerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrGl 701
 Qy 1182 -AGCTGCAACCCCAACATGCGTCCCGCCGCACAAAAGACAAAAGCGCTGACCGCGCAGCTCC 1240
 Db 701 uSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaPr 721
 Qy 1241 GTAACGAGCGCTGATGCCAAAACCTTCATTGTGTGGAGCGCAAAACGCATCTTTCATCGGCTCA 1300
 Db 721 o-----ValProThrProSer-----SerSerThrThrGl 731
 Qy 1301 TTCAACTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCGGTGTCATCGAAAGC 1360
 Db 731 uSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaPr 751
 Qy 1361 CCCAAATCCGAAACAGATGGAGCGCACCCCTCGCGATACCAACCCGAATACG----- 1415
 Db 751 oValThrSerSerThrThrGluSerSerSerAlaProValProThrProSerSerSerTh 771
 Qy 1416 -----CCTACCGGTTTACCTCGACAAACACACACC---GCCTGCAA 1453
 Db 771 rThrGluSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSe 791
 Qy 1454 TGGCAGCATCCCGCCACCCGAAAAAACCCTACCGAAACGAAC-----CGAA 1498
 Db 791 rAlaProValProThrProSerSerSerThrThrGluSerSerValAlaProValProTh 811
 Qy 1499 GCCA-----AAGTTTGGAAACGATCGCGCAAAATCC 1532
 Db 811 rProSerSerSerSerAsnIleThrSerSerAlaProSerSerThr 826
 RESULT 9
 US-09-252-991A-17112
 ; Sequence 17112, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUD
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 17112
 ; LENGTH: 655
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17112
 Alignment Scores:
 Pred. No.: 8,23e-16 Length: 655
 Score: 303.00 Matches: 155
 Percent Similarity: 33.5% Conservative: 37
 Best Local Similarity: 27.1% Mismatches: 179
 Query Match: 10.6% Indels: 202
 DB: 2 Gaps: 29
 US-10-665-990A-13 (1-1561) x US-09-252-991A-17112 (1-655)
 Qy 157 CCTGGACAACATCTCTGCAAT-----CCGGCACACCCCTCATATAACA 198
 Db 140 ProAlaGluGluProAlaGlyValArgGluLeuProArgHisProAlaGlyGly 159

```

QY 199 CGG-----GCTATCGACATCTACTGCTGACGACCCCGACGAGCCCTTGC 246
Db 160 ArgAspProAlaGlyArgLeuArgArgLeuValProArgArgProProArgHis----- 177
QY 247 CGCCCGCGCGCCCTTATCGAATCTGCGCAACACAGCCTCGATTTGCAATACTACTATTG 306
Db 178 -----TTPArgGlyLeuGlyArgValArgAlaValProArgHisAspLeuGlyArgAla 195
QY 307 GCGCAACGACATTTCCGCGAGCTGCTGTTCAACCTCATGTACTTGGCGGAGACGCGG 366
Db 196 AspGluAlaArgLeuArgArgCysProAlaProGlyLeuAlaProProArgArgThrLeu 215
QY 367 CGTGGCGTAGCTGCTGTTGGACCAACACGCGGGTGGACGATCTCCTGCT 426
Db 216 ArgArgArg-----ProLeuHisArgProGlnGlyAlaArgProGlyArgSerAlaGly 233
QY 427 CGCCCTCGACGCCATCCCATATCGA-----AGTGGCCCTGTTCAACCC 471
Db 234 AlaAlaGlyArgGlyLysAlaGlnArgLeuHisLeuGluProValAlaArgValArgArg 253
QY 472 CTTCGTCTACGCAATCGCGGCGACTCGGTACTGAC-----CGACTTCCCGCGCT 525
Db 254 LeuArgAlaValArgGlnValArgGlyAspValProGlyValArgArgProAlaAla 273
QY 526 CAACCGCGCGCATGCAACAACATCTTTACCGCGACAAACGCGCCACCATACTCGCGG 585
Db 274 GluProGluGluAla-----AspProGlyHis----- 282
QY 586 AGCAATATCGGCGACGAATACTTCAAAGTCGGTGAGGACACCGTTTTCGCGCACTGGA 645
Db 283 -----GlyHis-----ArgProGly 287
QY 646 CATCTCGCCACGCGCGCTGTCGCGAAGTATCGCACGACTTCGACCGCTACTGGGC 705
Db 288 ArgArgGlnArgGluValArgArgGlnProLeuSer-----ArgGlnThrAlaGly 305
QY 706 AAGCCATTCCGC-----CCACAACGC 726
Db 306 ArgAlaArgArgArgProAlaProAlaAspArgArgAlaGlyTrpGlnGlyProGlyArg 325
QY 727 CACGCGCATATCCGAGCGCAACATCGGCAAGGTCTTCAAGCACTCGGATACAACGA 786
Db 326 ArgArgHisProValValLeuHisLeuProGly-----LeuArg 339
QY 787 CGAAACATCCAGACGCGCTCCTCGCTACCGGAAACCGTCGAAAGTCCGCCCTCTA 846
Db 340 ArgGlyValProAspAspArgAlaArgArgHisArgArgHis---AlaProPro 358
QY 847 CCAAAAAATACAGACGGGACGCATCGACTGCGAGCGGT-----CCAAACCGC 894
Db 359 Pro-----AspProGlyGluArgArgHisAlaGluGlnGlyArgGlyValGlyGln 376
QY 895 CTTGATCAGGACACCCCTCGAAAGACTCGACG----- 930
Db 377 ProAspArgHisArgGlnProArgArgPheArgProArgProValGluLeuGlyGly 396
QY 931 CGACCGCGCAACCGCGATTGCGCGGAGGCTCGACAGCG----- 972
Db 397 ArgProGluProAlaAlaAspGluArgGluGluGluArgArgAlaValLeuGlyArg 416
QY 973 -----GTTCAACACGCGCAAAAAAGCGTCTA----- 999
Db 417 ArgTrpArgLeuArgHisAlaGlnProAlaHisProAlaArgLeuArgGluAspProGln 436
QY 1000 -----TCTGTTTACCCCTATTTCGTCCTTACAAAATCGCGCACAGCGCACTGGC 1050
Db 437 GlySerArgGlyGlyLeuArgArgAlaArgProGlyGlyAlaArgGlnArgArgArgGly 456
QY 1051 AAAAAGTGTGACGACG----- 1068
Db 457 ---AlaProAlaGlyArgArgGlyTyrLeuProGluProGlyGlnAlaGlnTyrArgAsp 475

```

```

QY 1069 -----CATAGAGTTTACCGTCTCGCTGACCACTCGCTACAGGC 1104
Db 476 ProGlyGluLeuProLeuHisProHisArgGlnLeuArgProAlaGlnLeuProCysAla 495
QY 1105 GACCGAGTTTCCCGCGCTCCATTCGGCTACGTCAAATACCGAAAAACCGCTGCTCAAAAGC 1164
Db 496 GluGluArgGlyLeuArgArgPro-----ArgArgArgLeuProGlyThrAlaProGlnHis 513
QY 1165 CGGCATCAAACTCTACGAGTGGCAACCAACCATCGCTCCCGCGCAAAAAGACAAAGG 1224
Db 514 LeuHisArg-----ArgThrAlaAlaArg-----ArgHisAlaAlaThrGly 527
QY 1225 CCTCACCGGACGCTCGCTAACACGCTGCATGCCAAAAACCTTCATTGTGGAGCGCAACG 1284
Db 528 ProGlnGln-----GlyArgGlnArg 534
QY 1285 CATCTTCATCGGTCTATTCAACCTCGACCCCGTTCCGCGACGCTCAATACCGAAATGGG 1344
Db 535 HisLeuProArgProLeuLeuPro----- 542
QY 1345 CGTCGTATCGAAGCGCCAAAATCGCAGAACACAGATGAGGCGCACCTCGCGATACCAC 1404
Db 543 -----GlyProLeuGlnArgArgValArgGlyAlaAlaGlnArgAlaAlaArg 558
QY 1405 ACCCGA-----ATACGCTTACCG----- 1422
Db 559 ThrArgHisArgGlyGlnGlyAsnAlaThrLeuArgLeuProLeuAlaLeuLeuArgArg 578
QY 1423 -----CGTTACCTCGCAAAACACACACCGCTGCAATGGCAGCA 1461
Db 579 ArgTrpArgAlaAspHisArgHisProArgGlnAlaThrAspSerArgHisAlaHis 598
QY 1462 TCCGCGCACCGAAAAACCTACCGAAGACCGCAACCGCAAACTTTGGAAACCGCATCGC 1521
Db 599 GlyArgHisProArgAsp-----Arg 605
QY 1522 CGCAAA-----AATCTATCTCTGCTGCTCCATCGA 1551
Db 606 ArgArgAlaGlyGlyArgArgLeuProAlaValHisArg 618

RESULT 10
US-09-252-991A-23696
; Sequence 23696, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23696
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23696

Alignment Scores:
Pred. No.: 9,67e-16 Length: 618
Score: 302.00 Matches: 173
Percent Similarity: 34.8% Conservative: 35
Best Local Similarity: 29.0% Mismatches: 193
Query Match: 10.6% Indels: 196
DB: 2 Gaps: 38

US-10-665-990A-13 (1-1561) x US-09-252-991A-23696 (1-618)
QY 169 CTGCAAAATCGGCGACACACCCCTCATAACAACGGGCTATCCGACATCTACCTGCTCGACGA 228

```

Db	308	GlnGlyAlaGluHisProGlyArgAlaAspArgLeuProAlaLeuArgProAlaAlaA	327	
Qy	1102	-----GGCGACCGACGTTCCCGCGCTTCATTCCGGCTACGTCAA	1149	
Db	328	ProAspLeuGlyGlnArgHisHisSerGlnPro--ArgValArgProGlyLeuProThr	346	
Qy	1150	ACCGTGCTCAAGCCGG-----CATCAAACTCTACGAGCTCAACCCAA	1194	
Db	347	AlaAlaAlaGluArgProLeuAlaGluArgLeuProAlaArgCysGlnTrpGlnTyr	366	
Qy	1195	-----CCATGTCGGCTCCCGCCACAAA	1215	
Db	367	ProAlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro	386	
Qy	1216	AGACAAAGGCGCTGACCGGACGTCGTGTAACCAACCGCTGCATGCCAA	1272	
Db	387	-----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuCysLeuArgArg	403	
Qy	1273	GGACGCCAAACGCAT-----CTTCATCGGCTCATTCAA	1305	
Db	404	GlyArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln	422	
Qy	1306	CCTCGACCCCGCTCCGCACGGCTCAATAC-----CGAAATGGCGCTCGTCATCGA	1356	
Db	423	ProGlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgArgAlaSerGly	442	
Qy	1357	AAGCCCCAAATCGCAGACAGATGAGCGCACCTCGCGCATACCAACCCCGAATACGC	1416	
Db	443	AspGlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArg	462	
Qy	1417	CTACCGGTTTACCCTCGACAAACACACCCCTGCAATGGCA-----	1458	
Db	463	ArgProGlyProAlaAlaArgGlnProArgProAlaGlyArgArgGlleProLeu	482	
Qy	1459	-----CGATCCCGCCACCGAAA-----AACCTACCC--	1485	
Db	483	GlnArgProArgProAlaHisArgGlnAlaHisProArgGluArgHisHisLeuProLeu	502	
Qy	1486	-----GAACGAAACCCGAAGCCAAACTTTGGAAACGCAT	1518	
Db	503	TrpSerArgArgProValAlaArgArgSerArgThrArgArgGlnArgProGluThrAla	522	
Qy	1519	CGCGCGAAAAATCTATC-----CTGTCTGCC-----CATCGAAGG	1554	
Db	523	SerProValLeuProValAlaGlyGlnProAlaAlaGlyHisHisArgArg	539	
RESULT 11				
US-09-252-991A-28358				
; Sequence 28358, Application US/09252991A				
; Patent No. 6551795				
; GENERAL INFORMATION:				
; APPLICANT: Marc J. Rubenfield et al.				
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUD				
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS				
; FILE REFERENCE: 107196.136				
; CURRENT APPLICATION NUMBER: US/09/252,991A				
; CURRENT FILING DATE: 1999-02-18				
; PRIOR APPLICATION NUMBER: US 60/074,788				
; PRIOR FILING DATE: 1998-02-18				
; PRIOR APPLICATION NUMBER: US 60/094,190				
; PRIOR FILING DATE: 1998-07-27				
; NUMBER OF SEQ ID NOS: 33142				
; SEQ ID NO 28358				
; LENGTH: 618				
; TYPE: PRT				
; ORGANISM: Pseudomonas aeruginosa				
US-09-252-991A-28358				
Alignment Scores:				
Pred. No.:		1.16e-15	Length:	618
Score:		301.00	Matches:	175
Percent Similarity:		34.6%	Conservative:	31
Best Local Similarity:		29.4%	Mismatches:	196

Query Match:	10.6%	Indels:	194
DB:	2	Gaps:	37
US-10-665-990A-13 (1-1561) x US-09-252-991A-28358 (1-618)			
QY	169	CCTGCAATCCGGCAGCACCCCTCATAAACAACGGGCTATCCGACATCTACCTGCTCGACGA	228
Db	4	ProAspHisProArgHisProTrp---ArgArgGlnLeuAlaGluLeuArgGlyLeuArg	22
QY	229	-----CCCCACCAAGCCCTTGCCGCCGCCGCCGCCCTTATCGAATCTGCCGAACA	279
Db	23	ArgThrAlaProGlyGlyAspArgGlnGlnProArgArgThrProArgIleArgArgArg	42
QY	280	CAGCTCGATTGCAATCTACTATTTGGCGCAACGACAT-	318
Db	43	-----HisGlnGlyGlnSerHisArgProAlaHisGlnGlyArg	55
QY	319	-----TTCGGCAGGCTGCTGTTCAACCTCAT	345
Db	56	GlnArgGlnProGlyAlaProAlaAlaMetGlyLeuArgArgAla-----TrpProAla	73
QY	346	GTACTTCCTCCGACAGCGCGTGCAGCTGCTGTGGACGACAAACAACACGCG	405
Db	74	AlaProCysGlyArgArgArgProAspThr-----GlnLeuArgLeu	88
QY	406	CGGTTGGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGCGCCTGTT	465
Db	89	ArgProGlnArgGlnSerGlyArgArgAsnGlnPro-----AlaProVal	103
QY	466	CAACCCCTT-----CGTCTTACGCAATGGCGCG	495
Db	104	ArgProGlnProGlyLeuArgArgProArgProAlaGlyArgProGluArgSerProArg	123
QY	496	ACTCGGCTACTGACCACTTCCCGCTCAACCGCGCATGCACAACA-	546
Db	124	ArgGlnAspProThrArgLeuArgArgProGlyGlnProHisArgGlyGlnGlyProAla	143
QY	547	-----ATCTTTACGCGCAGCAACCG-----	567
Db	144	ArgArgHisHisProLeuArgIleArgArgProArgGlnProAspProThrGlyGlnPro	163
QY	568	-----CGCCACCATCTCGCGCGCAACAATATCGGCGACGAATATCTCAAAGTCGGTGA	621
Db	164	GlyGlnArgHisHisLeuArgAlaArgArgArg-----GlnArgHisPro	180
QY	622	GGACACCGTTTTCCCGACTTGGACATCTCGCCACCGCGCAGCTGTCGCGCGAAGTATC	681
Db	181	ProHisArgArgSerArgCysGlyHis---ArgValSerLeuArgArgProGlnSerAla	199
QY	682	GCACGACTTCGACCGCTACTGGGCAAGCCATTCCGCCACACGCCACGCCGCATCATCCG	741
Db	200	GlyArg---AlaProLeuAlaGluArgPro---GluProArgArgThrVal-----Pro	215
QY	742	CAGCGCAACATCGGCAAGGTCTTCAAGCACTCGGTATACAACGAAACATCCAGACA	801
Db	216	LeuArgProHisArg-----ArgArgGlnProGlyHis	226
QY	802	CGGCTCTCGGCTACCGCAAAACCGTGAACAGTCGCCCTCTTACCAAAAATACAGAC	861
Db	227	ArgProProGly-----ArgHisArgArg-----CysProArgGlnProGly	240
QY	862	GGGACGATCGACTGGCAGAGCGTCCAAACCCGCCTGATCAGCAGCACCCCTGCAAAAGG	921
Db	241	ValProLeuArgArgAlaArgGlnProGlyArgAlaGlyThrGlnHisPro-----	257
QY	922	ACTCGACCGCGCAGCGCGCAAAACCGCGCATTTGCCGGGAG---GCTGCAAAACGCGCTCAA	978
Db	258	-----ProArgProAlaAspProAlaArgProArgAspLeuProLeuArgArgGlyGln	275
QY	979	ACAGCCGAAAAAGCGTCTATCTGGTTTCAACCTATTTCGTCCTCACAAATCCGGCAC	1038
Db	276	ProThrAlaGlyAspArgLeu-----ProLeuArgProArgHis	288

QY 1039 -----AGACGCACTGGCAAA-----ACTGGTCAGGACGGCAT 1071
Db 289 ArgLeuProAlaGlnArgArgProGlyArgGlnArgAspProGlySerGlyArgGln 308
QY 1072 AGCGTTACCGTCT-----GACCAACTCGCTACA----- 1101
Db 309 GlyAlaGluHisProGlyTyrAlaAspArgLeuProAlaLeuArgProAlaAlaPro 328
QY 1102 -----GGCGACCGACGCTTCCGCGCTCAATTCGGCTACGTCAA---ATACCGAAAAACC 1152
Db 329 AspLeuGlyGlnArgHisHisSerGlnPro---ArgValArgProGlyLeuProThrAla 347
QY 1153 GCTGCTCAAGCCGG-----CATCAAACTCTACGAGCTGCAACCCAA----- 1194
Db 348 AlaAlaGluArgProLeuAlaGluArgLeuProAlaArgCysGlnTyrPro 367
QY 1195 -----CCATGCGCTCCCGCCCAAAAAA 1218
Db 368 AlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro--- 386
QY 1219 CAAAGGCTCACCAGGAGCTCCGCTAACCGAGCTCGATGCCAA---AACCTTCATTGTGGA 1275
Db 387 -----ProaspArgGlyThrArgArgProGlyArgAlaGlnLeuArgLeuArgArgGly 404
QY 1276 CGGCAAAACGGAT-----CTTCATCGGCTCATCTAACCT 1308
Db 405 ArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---GlnPro 423
QY 1309 CGACCCCGCTTCGCGACGGCTCAATAC-----CGAAATGGGCGTGTCTACGAAAG 1359
Db 424 GlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgAlaSerGlyAsp 443
QY 1360 CCCCAAAATCGCAGAACAGATGAGCGCACCCTCGCGATACCACACCCGAAATACGCCTA 1419
Db 444 GlnArgArgArgGlnProHisProGlyProArgAlaGlnThrGlyLeuArgArg 463
QY 1420 CCGCGTTACCTCGCAACACACAAACCGCTCGCAATGGCA----- 1458
Db 464 ProGlyProAlaLysArgGlnProArgArgProAlaGlyArgArgIleProLeuGln 483
QY 1459 -----CGA 1461
Db 484 ArgProArgProAlaHisArgGlnAlaHisProArgGluHisHisLeuProLeuArg 503
QY 1462 TCCGCGCACCGAAAAACCTACC CGAACGACCCGAAGCCAACTTTGGAAACGCATCGC 1521
Db 504 SerArgProValAlaArgArgSerArgSerArgGlnArgProGluThrAlaSer 523
QY 1522 CGCAAAAATCCTATC-----CCTGCTGCC-----CATCGAAGG 1554
Db 524 ProValLeuProValAlaGlyGlnProProAlaAlaGlyHisHisArgArg 539
RESULT 12
US-09-252-991A-23598
; Sequence 23598, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE OF INVENTION: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23598
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23598

APPLICANT: Marc J. Rudenfeld et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

RESULT 12
US-09-252-991A-23598
; Sequence 23598, Application US/09252991A

; GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.

1. TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID ANALYSIS

;; TITLE OF INVENTION: AERUGINOSA FOR DIAG
;; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A
 : CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 23598

```

; LENGTH: 507
; TYPE: PRT

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23598

Alignment Scores:

Pred. No.: 3,25e-15 Length: 507
Score: 295.00 Matches: 149
Percent Similarity: 36.8% Conservative: 32
Best Local Similarity: 30.3% Mismatches: 159
Query Match: 10.3% Indels: 112
DB: 2 Gaps: 27

US-10-665-990A-13 (1-1561) x US-09-252-991A-23598 (1-507)

QY 196 CAACGGCTATCCGACATCTACCTGCTCGACACCCCAAGAGCCCTTGCCTCGCG--- 252
DB 11 GlnGlnArgLeuArgHis-----AlaArgGlnProGlyArgHisGlnProAlaProGly 28
QY 253 ---CGCGCCCTTATCGAATCTGCCGAACACACAGCCTCGATTTGCAATCTACATTTGGCG 309
DB 29 ArgArgArgPro----- 32
QY 310 CAACGACATTTCCGGAGGCTGCTGTTCAACCTCATGTA---CCTTGGCGGAGACGGCG 366
DB 33 -----AlaGlnProAlaGlyArgProAspLeuGlnProGlyArgArgProArg 48
QY 367 CGTGGCGGTACGCTGCTGTGGAGACAAACACAGCGCGGTGGACGATCTCTGCT 426
DB 49 GlyAla-----GlyArgGlnProGlyGlnProGlyHisAlaProGly 63
QY 427 CGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTGTCTCTACGCAA 486
DB 64 ProGlyArgGlnProAlaAlaTyrArgArgAlaGlyGlnProGlyArgProGly 83
QY 487 ATGGCGCGCACTCGGCTACTGACCGACTTCCCGCGCTCAACCGCGCATCCACAA 546
DB 84 Gln-----ProGlyArgGlnProArgProAla----- 92
QY 547 ATCCTTTACCGCGCAACCGCGCACCATCTACGCGGACGCAATATCGCGAGCAATA 606
DB 93 -----GluArgGlnProArgGlnArgArgArgAlaGlnGlnArgGlnGly 109
QY 607 CTTCAAAGTCGTGTGGACACACGTTT-----CGCGACCTGGACAT 648
DB 110 LeuAlaGluAla---GlyHisArgAlaValArgGlnArgArgHisProGlyAla 128
QY 649 CTTCCGACCGGAGCGTGTGGGAGTATCGACAGCTTCGACCGCTACTGGGCAAG 708
DB 129 ValAlaGlyAspSerArgArgAlaArgAlaGlnProAlaGlyProSerLeuGlyAla 148
QY 709 CCATTCGCGCCCAACCGCACGCGCATCATCCGACGGCAACATCGGCAAGGCTTCA 768
DB 149 Gly---ArgArgGlnProHisArgHisArgArgLeuArgGln-----ProGly 163
QY 769 AGCACTCGGATACACGACGAAACATCCAGACCGCTCTCGCTACCGGAAACCGT 828
DB 164 ArgArgProLeuArgGlnArgProAlaGlnProArgArgProAlaLeuProGlnProGly 183
QY 829 CGAACAGTCCGCTCTACCAAAATACAGCGGACCGCATCGCTG-----GCAGAG 882
DB 184 ArgGlyGlyProGlyArgGlnGlyArgArgArgAlaHisArgLeuGlnProGlyArg 203
QY 883 CGTCCAAACCGCGCTGATCAGGACACCCCTCGCAAAAGGACTCGACCGCGACCGCAA 942
DB 204 ArgAlaGlyGlnPro---LeuArgProValGlyLysArgLysArgAlaProAlaArg 222
QY 943 ACCCGCATTTGGGAGGCTGCAAGACGCGTCAACAGCCCGCAAAAGCGTCTATCT 1002
DB 223 ArgArgAsp---ArgGlnGlnArgArgGlnProAlaArgProArgProGlnArgGlnLeu 241
QY 1003 GGTTCACCTATTTCTCTCCCTACAAATCCGGCACAGCGACTGGCAAACTGGTGA 1062
DB 242 AlaValGlyArgTrpArgProGluGln---ArgLeuArgArgAlaGlyLysArgGlnPro 260
QY 1063 GGACGGCATAGAGTTACCGTCTCTGACCAACTCGCTACAGGGGAC-----CGA 1110

Db 261 GlyPro-----ArgPro-----AlaThrGlyGlnProGlyGlnArgArg 273
QY 1111 CTTTCCCGCGCTCCATTCGGCTACGTCAATA-----CCG 1146
DB 274 ArgAlaHisProProTyArgGlnArgHisLeuArgProGlyPheArgAlaGlyAspPro 293
QY 1147 AAACCC-----GCTGCTCAAAGCGGCAATCAA 1173
DB 294 ArgArgArgArgThrAspHisGlnTrpProAlaGlyHisProArgGlnArgMetAspGln 313
QY 1174 ACTTACGAGCTGCAACCCCAACCATGCGCTCCCGCCCAAAAGACAAAGCGCTCACCGG 1233
DB 314 GlnGlnArgAlaAlaSerArgThrProGluProGlyHisArgHisLeuProProAspGly 333
QY 1234 CAGCTCCGTAAACCGACCTGCA---TGCCAAACCTTCTTGTGGACGGCAACGCAATCTT 1290
DB 334 -----ArgGlyGlnAlaAlaGlyAlaValLeuHisTrpProArgArg----- 348
QY 1291 CATCGGCTCATCAACCTCGACCCCGCTTCGCGACGGCTCAATCCGAAATGGCGTCT 1350
DB 349 ---ArgLeuGlnArgArgProAlaGlyGlnArgArgGlnLeuAlaThrArgAlaGlu 367
QY 1351 CATCAAAAGCCCAAAATCGCAGAACAGATGAGCGCACCTCGCGCATACCCACCCCA 1410
DB 368 ArgArgLeuProTrpGlnArgProArgHis-----GlnProArgArgLeuArgProGlu 385
QY 1411 ATACGCTTACCGGTTTACCTCGACAAACACACACCGCTCGCATGGCAGATCC----- 1464
DB 386 ArgArgGlnProArg---ProArgGlnCysArgGlnProArgArgArgGlnCysHis 404
QY 1465 -----CGCCACCGCGAAACCTACCC-----GAACGAACCCGA 1497
DB 405 AlaTrpArgArgGlnProAlaGlyGlnProTrpAlaAspHisArgCysArgArgProArg 424
QY 1498 AGCCAAACTTTGGAAACGATCGCGCGCAAAATCTCT 1533
DB 425 GlyGlnArgArgGluProGluGlnLeuArgHisPro 436

RESULT 13

US-09-252-991A-16789
; Sequence 16789, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16789
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16789

Alignment Scores:

Pred. No.: 5,52e-15 Length: 774
Score: 293.00 Matches: 163
Percent Similarity: 34.9% Conservative: 41
Best Local Similarity: 27.9% Mismatches: 229
Query Match: 10.3% Indels: 152
DB: 2 Gaps: 33

US-10-665-990A-13 (1-1561) x US-09-252-991A-16789 (1-774)

QY 109 GGAAGAACGCGAAGCCGTCATTTCAATCTTCCAAACCTGCTCTCTCGACAACT 168
DB 58 GlyArgGlnAsp---LeuProLeuLeuGln-----ProProGlySerArg 71

US-10-665-990A-13 (1-1561) x US-09-134-001C-4273 (1-495)

```
QY 191 CATAACAACGGGCTA-----TCCGACATCTACCTGCTCGACAGCCCCAC 235
Db :|||||
122 TyrAsnAsnAlaAlaPheLeuThrThrAspAsnAspLeuThrIleTyrThrAspGlyHis 141
QY 236 GAAGCCCTTCCGCGCGCCGCTTATCGAATCTGCGGAACACAGCTCGATTGCAA 295
Db :|||
142 GlnLysPheAspAspLeuileAsnAspIleArgHisAlaGlnSerTyrIleHisIleGln 161
QY 296 TACTACATTTGGCGGAACACATTTCCGGCAGCGCTGTTCAACTCATGTACCTTGGC 355
Db :|||||
162 TyrTyrIleIleHisSerAspAsnLeuGlyLysGlnLeuLeuHisIleLuleuLysLys 181
QY 356 GCAGAACGCGGGTGGCGCTAGCCCTGCTGTTGGACGACAAACACACGCGCGTTGGAC 415
Db :|||||
182 AlaGluGluGlyIleGluValLysMetLeuTyrAspAspMetGlySerArgAspLeu--- 200
QY 416 GATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCTGTTCAACCCCTTC 475
Db 201 -----ArgLysLys 203
QY 476 GTCCTACGCAATGGCGGCACCTCGGCTACCTGACCGAC-----TTCCCC----- 520
Db :|||||
204 AspLeuLysLysPheArgGlnLysGlyHisAlaGluSerPhePheProSerLysLeu 223
QY 521 -----CGCCTCAACCGCGCATGCACAAACAAATCTTTTACCCCGCAGCAAC 565
Db :|||||
224 ProLeuileAsnLeuArgMetAsnAsnArgAsnHisArgLysIleValIleAspGly 243
QY 566 CGGCGCACCATCTCGCGGCGCATATCGGGCAGCAATCTTCAAGTCGCTGAGGAC 625
Db :|||||
244 ThrIleGlyTyrValGlyGlyPheAsnValGlyAspGluTyr-----IleGlyLysSer 261
QY 626 ACCGTTTTCGCC-----GACCTGGACATCTCGCCACCGGACGCGCTCGCGCGAA 676
Db :|||||
262 LysLysPheGlyTyrTrpArgAspThrHisLeuArgIleLysGlyAspAlaValAsnAla 281
QY 677 GTATCGCACACTTGCACCGCTACTGGGCAAGCCATTCGCCCCAC---AACGCCACGCGC 733
Db :|||
282 LeuGlnLeuArgPheIleLeuAspTrpAsnSerGlnSerThrArgAspAsnLeuThrTyr 301
QY 734 ATCATCCGC-----ACGGCAACATCGGCNAGGCTCTTCAAGCA 772
Db :|||
302 GluSerArgTyrPheProAspValAspSerGlyThrIleGlyIleGlnIleAlaSer 321
QY 773 CTCGGATACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCGAA 832
Db :|||
322 SerGlyProAspGlu----- 326
QY 833 CAGTCGCCCTCTACCAAAAAATACAGACGGGACGATCGACTGGCAGAGCGCTCAA--- 889
Db :|||||
327 -----AspTrpGluGlnIleLysTyr 333
QY 890 -----ACCGGCTGATCAGCAGACCCCTGCAAAAGGACTCGACCGCGCCGCAAA 943
Db :|||||
334 GlyTyrLeuLysMetIleSerAla----- 342
QY 944 CGCGGATTGCGGGAGGCTGCAAGACGCGCTCAAAACGCGCTCAACAGCCGAGCGCTCTATCG 1003
Db :|||||
343 -----LysGluSerIleTyrIle 348
QY 1004 GTTTCACCTATTCTCTCTCAAAAATCCGGCAGACAGCAGCTGTGCAAAACTGTGTGACG 1063
Db :|||||
349 GlnSerProTyrPheIleProAspGlnAlaPheLeuAspSerIleLysIleAlaAlaLeu 368
QY 1064 GACGGCATAGCTTACCGTCTGACCAACTCGCTACAGGCGACGACGCTTGGCGCCGNC 1123
Db :|||||
369 GlyGlyValAspValAsnIleMetValProAsnLysArg-----AspHisProPheVal 386
QY 1124 CATTCGGGTACTGCTCAAAATACCGAAAACCGCTGCTCAAAAGCGGCACTCAAACTCTACGAG 1183
Db :|||
```

```
Db 387 TyrTrpAlaThrLeuLysAsnValAlaSerLeuLeuGluAlaGlyValAsnValTyrHis 406
QY 1184 CTGCAACCAACCATGCGCTCCCGCCACAAAAGAGCAAAAGCGCTACCGGAGCTCCGTA 1243
Db :|||
407 Tyr-----AspAsnGlyPhe----- 411
QY 1244 ACCAGCCTGCGATCCAAAACCTTCAATTGTGGACGGCAACGCATCTTTCATCGGCTCATTC 1303
Db :|||||
412 -----LeuHisSerLysThrLeuValIleAspAspGluValAlaSerValGlyThrAla 429
QY 1304 AACCTCGACCCCGCTTCGCGCAGGCTCAATACCGAAATCGGCGTCTCATCGAAAGCCCC 1363
Db :|||||
430 AsnMetAspAsnArgSerPheThrLeuAsnPheGluValAsnAlaPheIleTyrAspGlu 449
QY 1364 AAATCGCAGAACAGATGAGCGCACCTCTCGCCGAT----- 1399
Db :|||||
450 GlyValAlaArgSerLeuLysGlnAlaPheIleAsnAspMetLysLeuSerAsnLysLeu 469
QY 1400 ACCACACCGGATAGCGCTACCGGCTTACCTCGACAAA 1438
Db :|||||
470 ThrSerGluGluTyrAlaLysArgAsnLeuValLys 482

RESULT 15
US-09-252-991A-21881
; Sequence 21881, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21881
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21881

Alignment Scores:
Pred. No.: 1,028-14 Length: 720
Score: 289.50 Matches: 160
Percent Similarity: 34.1% Conservative: 41
Best Local Similarity: 27.2% Mismatches: 209
Query Match: 10.2% Indels: 179
DB: 30 Gaps: 30

US-10-665-990A-13 (1-1561) x US-09-252-991A-21881 (1-720)
QY 112 AGAAGCGGACGGAAGCGCTCAATTCATTAATCTTCCAAACCTGCTCTCGCAACATCCT 171
Db :|||
179 ArgAlaThrGlyGlnProArgArgGlnGlyHisArgArgGlnArgProAlaProHisArg 198
QY 172 GCAAATCCGGCACACCCCTCATAAACACGGGGTATCCGACATCTACCTGCTCGACGCC 231
Db :|||
199 AlaAlaProAlaGluProGlyGlnGlyAlaGlyArg-----AlaArgArgPro 215
QY 232 CCACGAAGCCTTGGCGCGCGCGCTTATCGAATCTGCGAACACAGCCTCGATT 291
Db :|||
216 -----GluProGlyArgArgArg----- 221
QY 292 GCATATACATTTGGCGCAACGACATTTCCGCGAGGCTGCTTTCACCTCATGTACCT 351
Db :|||||
222 -----ThrValGlnArgArg---ArgArgSerAlaArgGlnProGluGlnPro 236
QY 352 TCCCGCAGAACCGCGGTGCGGTGCGCTGCTGTGTGGACGACAAACAACCGCGGGT 411
Db :|||
237 GluArgGluProArgArg-----ArgAlaGlyGlnPro 247
```



```
Db 71 SerLeuPro-----ThrArgProArgHisArgAlaProGly 82
|||
577 -----ACTCGGGACGCAATATCGG---CGACGAATA----- 606
:|:|:|
83 HisGlyGluProSerArgProAlaProAlaArgCysArgArgAlaAlaValaspProGly 102
|||
607 -----CTTCAAAAGTCGTGAGGACACCGTTTTCGC 636
:|:|:|
103 AspAspGlyGlyAsnGlyAspLeuAlaGlyIleGlnProArg-----ProPheArg 119
|||
637 CGACTTGA-----CATCTCGCCACGGCAGCGTCTCGG 672
|||
120 ArgHisGlyProHisProAlaLeuGlyGlnProAspProHisHisAlaAspArg----- 137
|||
673 CGAAGTATCGCAGCTTCGACCGGTACTGGGCAAGCCATTCGGCCCAACGCCACGCGG 732
:|:|
138 -----LeuLeuLeuLeuArgProValLeu-----ProTrpArgProAlaArgProAla 153
|||
733 CATCATCCGACGGCAACATCGGCAAGGG-----TCTTCAAGCACTCGGTATACAAGA 786
|||
154 HisProProAspHisGlyArgLeuGlyValAlaGlyAspArgArgArgLeuArgArg 173
|||
787 CGAAAC---ATCCAGACCGCCTCT-----GCCTACCGGAAACCGTCGACAGTC 837
|||
174 ArgHisLeuValAspHisArgProGlyArgAlaValPheArgArgGlyHisVal 193
|||
838 GCCCTCTACCAAAAAATACAGAC-----GGGACGATCGACTG 876
|||
194 ArgProValProArgArgProLeuProArgThrProArgAlaGlyAlaHisArgArg 213
|||
877 GCAGAGCGT-----CCAAACCGC----- 894
|||
214 LeuHisArgAlaThrGlyGlnProAlaAlaGlyValLeuProAlaProGlyArgArgGln 233
|||
895 CTGTATCAGCGACACCCCTGCAAAAGGACTCCAGCGACCGCCGCAACCCCGATTGC 954
|||
234 ProGlyArgAlaHisProAlaGlnArgThrAlaArgArgProArgThrGlyThrAla 253
|||
955 CGGGAG-----GCTGCAAGCGCGCTCAACAGGCCGA-----AAAAAGGCTCTA 999
|||
254 ArgArgAlaAlaGlyArgArgSerAspProArgArgProValGlnArgArgVal 273
|||
1000 TCTGTTTCCACCTATTCTGCTCCCTACAAATCCGGCAGACGCACT---GGCAAAAT 1056
|||
274 AlaAla-----HisArgArgValProAlaProSer 283
|||
1057 GGTGACGAGCGCATAGACTTTACCGT-----CCTGACCAACTCGTACAGCGACCGA 1110
|||
284 ThrProGlyArgArgArgHisArgArgHisProGluArgArgGlyAlaAlaAspArg 303
|||
1111 CGTTGCGCGCTCCATTTCGGGTACTGCAAAATCCGAAACCGCTGCTCAAAAGCGCGCAT 1170
|||
304 ArgGlyArgArgProTrpArg---ArgGlnProPro-----ValGlyHis 317
|||
1171 CAAACTCTACGAGCTGCACCCCAACCATCGCTCCCGCCCAAAAGACAAAGCGCTGAC 1230
|||
318 ArgPro---ProAlaGlyThrArgPro-----GlyGlyGlnAlaAlaProGly 332
|||
1231 CGGACGCTCGTAACCGACTCATGCTCCCAAAACCTTTAT-----TGTGGACGGAAACG 1284
|||
333 GlyThrGlyArgGlnGlyGlyAlaValValProAlaHisArgProAlaGlyArgArgSer 352
|||
1285 CATCTTTCATCGGCTCATCAACCTCGACCCCGGTTCCGACCGCT----- 1329
|||
353 GlyArgArgAlaValValAlaGlyArgProProAlaArgLeuLeuGlyGlyAlaValAla 372
|||
1330 -----CAATACCGAAATGGGCTGCT 1350
|||
373 AlaGlyArgHisLeuProLeuArgProValAlaGlyHisAlaHisArgAlaHisArgGly 392
|||
1351 CATCGAAAGCCC---CAAAATCGCAGAACGATGGAGCGCACCTCGC----- 1395
|||
```

```
Db 393 HisArgGlnProAlaGlnAlaArgProAlaAspHisAlaArgAlaArgAlaGlyArgAla 412
1396 -----CGATACCAACCCGAATACGCTACCGCTTACCTCGCAACAAACACCGCCT 1449
|||
413 GluProHisArgHisGlyAsnLeuArgGlnAspArgHisProHisArgArgProAsp 432
|||
1450 GCAATGGCAGATCCCGCCACCGCAAAACCTACCCGAACGAAACCGAAGCCAACTTTG 1509
|||
433 AlaGlnAlaGlyAla-----ThrProAlaArgSerArgArgGlySerLeu 447
|||
1510 GAAACGCATCGCGC-----AAAAATCTATCTCTGCTGCGCCATCGA 1551
|||
448 ProGlyProArgArgArgProArgGluProPheArgThrProHisArg 463

RESULT 17
US-09-252-991A-31760
; Sequence 31760, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31760
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31760

Alignment Scores:
Pred. No.: 2,26e-14 Length: 1706
Score: 287.00 Matches: 163
Percent Similarity: 32.1% Conservative: 52
Best Local Similarity: 24.3% Mismatches: 207
Query Match: 10.1% Indels: 248
DB: 2 Gaps: 31

US-10-665-990A-13 (1-1561) x US-09-252-991A-31760 (1-1706)
QY 199 CGGGCTATCCGACATCTACCTCTCGACGACCCCAAGCCCTTGGCGCCGCGCGC 258
|||
998 ArgLeuLeuHisLeuArgProGlnArgProGlyGlnLeuProAlaGlnLeuArgArg 1017
|||
259 CCTTATCGAATCTGCGGAACACAGCTCGATTGCAATATCATTTGGCGCAACGACAT 318
|||
1018 AlaArgAlaAspArgArgAlaPro-----ProGlnArgGly 1030
|||
319 TTCCGCGAGGCTGCTTCAA-----CCTCATGTACCTTGC-----CGCAGAAG 363
|||
1031 GlyAspProArgAlaAlaGlnGlyGlyAspHisLeuAspAlaAspArgHisLeu 1050
|||
364 CGGCGTGGCGGTACCGCTGCTTGGACGACCAACACACGCGCGGTGTCAGCTCTCT 423
|||
1051 ArgHisAlaArgHisProAlaLeuAlaGluGlnProAlaArgLeuProAlaArg 1070
|||
424 GCT---CGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCTCT 480
|||
1071 GlyIleArgProArgThrValArgArg-----ArgValArgArgTrpArgPro 1086
|||
481 ACGCAATGGCGCGCACTCGGTACTGTACCGACTTCCCGCCCTCAACCGCGCATGCA 540
|||
1087 ThrAlaArgAlaProAlaHisArgProGlnArgProAlaProProAlaLeuPro---Ala 1105
|||
541 CAACAAATCTTATCCGCGCAGCAACCGCCACCAT-----ACTCGCGCG 585
|||
1106 ThrArgGluGlnArgAspArgArgValHisProLeuAlaGlyLeuValArgArg 1125
|||
```

```

QY 586 ACGCAATATCGCGACGAATACTTCAAAGTCGGTGAGGACACCGTTTTTCGCGCGACTCGGA 645
Db 1126 ThrArgArgArgArgArgArgLeuArgProArgLeuArgHis-----ArgLupProGly 1143
QY 646 CATCCT-----CGCCACCGCAGCGTCGTCGGCGA----- 675
Db 1144 HisArgProLeuArgHisArgHisArgMetArgArgProLeuProProAlaAlaGlyThr 1163
QY 676 -----AGTATCGCAGCACTTCGACCGCTACTGGGCAAGCCATTTCGGCCCA 720
Db 1164 ArgProArgProGlyAsnLeuAlaThrLeuGlyAlaAlaGlyAspProAlaProPro 1183
QY 721 -----CAACGCCACCGCATCATCCGCGAGCGG 747
Db 1184 ProGlyPheLeuProGlyLeuValProArgArgArgGlnArgThrArgThrAlaAlaArg 1203
QY 748 CAACATCGGCAAGGCTCTTCAAGCACTCGG----- 777
Db 1204 ArgHisArgGluGlyProGlyAlaGlnArgGlyAsnProSerHisGlySerSerGlyGly 1223
QY 778 -----ATACAACGA----- 786
Db 1224 LeuProMetSerGlyProLysValValArgIleValThrArgGluGluAlaIleAlaThr 1243
QY 787 -----CGAAACATCCAGCACACGCGCTCCTCGCTTACCGCGAAACCGTCGGAACAGTCGCC 840
Db 1244 CysGluArgAspLeuGlnArgLeuAspLysAlaLeuAlaArgTrpGluAsnGlnAlaSer 1263
QY 841 CCTTACCAAAATAATACAGAC-----GGGACG 867
Db 1264 ArgLeuAlaGlnLeuSerAspAlaGluArgAlaAlaHisAlaArgArgAlaSerLeu 1283
QY 868 CATCGACTGCGCAGAG----- 882
Db 1284 HisAlaLeuAlaArgThrGlyThrLeuAlaGlyArgAlaThrThrGlyGluAspArgLys 1303
QY 882 ----- 882
Db 1304 ArgIleProGlnAlaArgProArgArgThrArgArgAlaArgHisProProGlyArgGly 1323
QY 883 -----CCTCCAAACCCCGTGAATCAGCGACACCCCTCCAAAGAGGACTCGACCG 930
Db 1324 IleThrProAlaThrProProProGlyLysArg-----GlnArgProAlaPro 1340
QY 931 CGACGCGCGCAACCGCGGATTCGCGGAG-----GCT 963
Db 1341 GlyProArgArgGlnAlaArgCysArgGlnArgGlySerProAlaAsnProAlaSerAla 1360
QY 964 GCAAGACGCGCTCAACACAGCCGCAAAAGGCTCTATCTGGTTTCACCCCTATTTTCGTCGC 1023
Db 1361 GlyArgArgArgProProArgArgArgGlySerProAlaArgProArgPheCysArgPro 1380
QY 1024 TACAAATCCGGCACAGA----- 1041
Db 1381 ArgGlnArgProArgArgGlyThrProGlnArgSerProAlaArgThrArgProAlaPro 1400
QY 1042 -----CGC 1044
Db 1401 GluAspArgArgAsnAlaHisHisProGlyThrValAlaArgProProAlaThrGlyArg 1420
QY 1045 ACTGGCAAAACTGTGCGAGCAGCGCATAGA---CGTTACCGCTCCTGCAACCTCGCTACA 1101
Db 1421 ProAlaArgThrAlaProGlyAlaHisArgProAlaHisArg-----ArgThrAlaThr 1438
QY 1102 GCGACCGCACTTCGCGCGTCCATTCGGCTACGTCAATACCGAAACCGCTGCTCAA 1161
Db 1439 AlaAlaArgArgGlyGlnArgPro-----SerValProArgAlaThrGlyThr 1454
QY 1162 AGCCGGCATCAACTCTACGAGCTGCAACCCAA----- 1194
Db 1455 ArgArgSerArgThrAlaProGlyAlaProGlnProAlaAlaArgGlnProGlyProArg 1474

```

```

QY 1195 -----CCATGTCGCGT 1203
Db 1475 ProArgProGlyArgProArgThrProAlaAlaThrProAlaProGlyThrProAlaArg 1494
QY 1204 CCCCGCCACAAAGACAAAGCCCTGACCGGAG-----CTCCGTAAACAGCCTGCTGATGC 1257
Db 1495 ProArgGlnArgSerArgArgProArgProArgArgArgThrArgArgThrAlaProAlaArg 1514
QY 1258 CAAACACCTT-----CATTGTGGACGGCAA 1281
Db 1515 ArgArgLeuProAlaArgGlnArgSerAlaAlaThrArgArgThrHisArgAlaMetGln 1534
QY 1282 ACGCATCTTCATCGG-----CTCATTCAACTCGACCCCG----- 1317
Db 1535 ArgHisProHisArgProProProProAlaThrGlyArgProArgProProGlyArgAla 1554
QY 1318 -----TTCGCGACGGCTCAATACCGAAATGGCGTCTCATCGCAAGCCCAAAATCGC 1371
Db 1555 ProGlyProArgGlnProArgLeuArgSerProArgArgHisGlyHisArgLeuGlyAla 1574
QY 1372 AGAACATGAGGCGCACCTCGCGGATACACACCGCAATACCGCTACCGCGTTACCCCT 1431
Db 1575 AspArgGlnGlyArgProAlaGlnThrArgHisProArgLeuArg---ProGlySerArg 1593
QY 1432 CGACAA-----ACACAAACCGCTGCAATGCGACGATCCGCGCACCCGAAAAACCTACCC 1485
Db 1594 ArgGlnGlyArgGlnArgProProAlaThrAlaArgArgArg-----Pro 1608
QY 1486 GAACGAAACCGAACCTTTGGAAACGCATCGCGCGCAAAATCTCTATC----- 1536
Db 1609 GlnArgGlnProArgGlnProAlaArgPro---ArgHisArgAsnProLeuValArgArg 1627
QY 1537 -----CCTGCTGCGCCATCGAAGG 1554
Db 1628 ValProThrProProGlyProAlaArgArg 1637

RESULT 18
US-09-828-523A-22
; Sequence 22: Application US/09828523A
; Patent No. 6764823
; GENERAL INFORMATION:
; APPLICANT: The Pharmacia & Upjohn Company
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 268 62120101
; CURRENT APPLICATION NUMBER: US/09/828,523A
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/266,327
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-828-523A-22

Alignment Scores:
Pred. No.: 1,53e-14 Length: 493
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x US-09-828-523A-22 (1-493)
QY 110 GAAGACGCGACGAAAGCCGTCATTTCAATATCTCCAAACCTGTCCTCTGGAC----- 163
Db 104 AspAsnLysGlnValGlnLysHisHisAspLeuValArgMetLeuLeuMetAspGlnAsp 123
QY 164 -----AACATCTCGCAATCCGCGCACACCCCTCATCAACACGCGGTATCCGAC 211
Db 124 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyArg 143

```



```
Db 134 SerThrAlaThrAla-----ThrArgThrTrpProThrAlaSerArg 147
QY 522 GCGTCAAC-----GCGCATGCAACAACAAATCTTTACCGCGCACAAACCGCGCCACCA 575
Db 148 AlaSerThrProGlyAlaCysSerThrGlySerProProSerAsnCysAlaArgPro 167
QY 576 TACTCGGCGAGCGCAATA---TCGGGAGCAATATCTTAAAGTCGGTGAGGACACCGTTT 632
Db 168 SerAlaArgProSerThrSerSerThrLysArgSerAlaSerAlaArg-HisArg-- 186
QY 633 TCGCCGACCTGCACATCTCGCCACCGCAGCGTGTGGCGAAGTATCGCAGCACTTCG 692
Db 187 ArgAspProGlyArgProAlaHisArgArg----- 197
QY 693 ACCGCTACTGGGCAAGCCATTCCGCCCAACCGCACCGCATCATCGCAGCGCAACA 752
Db 198 -----HisLeuProArgHisPr 203
QY 753 TCGGCAAGGCTTCAAGCACTCGGATACACGACGAAAC----- 792
Db 203 oArgArgGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 223
QY 793 -ATCCAGACAGC-----GCTCTCGCTACCGCGAAA 824
Db 223 uGlnArgArgProGlyGlnHisGlyProProProLeuAlaProGlyAlaProGlyVa 243
QY 825 CGCTCGAACAGTCGCCCTCTACCAAAA-----AATACAGACGG 863
Db 243 lLeuLeuGluAspAlaProLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 263
QY 864 GAGCATCACTG-----GCAGAGGTCCA-----AACCGCTGATCAGCGACA 908
Db 263 yArgHisArgLeuAlaGlyGlnAlaAspArgProArgThrArgProAlaAspAlaAla 283
QY 909 CCCTGCAAAAGACTCGACCGCGA-----CGCGCGCAAAACCGCGATTG 953
Db 283 pProGlyArgArgProArgGlyArgProAlaArgLeuGlyProProCysLeuAlaAspPr 303
QY 954 CGGGAG-----GCTGCAAGACGGCTCAACAGCCCGCGCAAAAGCGTCT 998
Db 303 oArgArgArgLeuGluArgAlaAlaAlaAlaAlaAlaAlaAspProProArgGlnArgLeuPr 323
QY 999 ATCTGGTTTC-----ACCTATTTCGTCTACAAAATCCGGCAGCAGG 1043
Db 323 oGlyGlyHisProArgProAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 343
QY 1044 CACTGCAAAACTGGTGACGAGCGCATAGAGTTACCGTCTGACCAACTCGCTACAGG 1103
Db 343 gProGly-----GlyArg-----Gly 348
QY 1104 CGACCGACGTTCCGCGCTCAATCCGGCTACGTCAATACCGAAACCGCTGCTCAAG 1163
Db 348 yGlnAlaAlaLeuArgLeuProLeuGlnArg-----LeuProArgThr-----Ar 364
QY 1164 CGGGATCAAACTCTACGAGCTGCAACCAACCATGCTCCCGCCCAAAAGCAAG 1223
Db 364 gArgValProProAlaProGlyArgArgGlnProTrpHisProAlaGlyAlaGlyGlnPr 384
QY 1224 GCGTGAC-----CGGCAGCTCGTAACCAAGCTGCATGC-----CAAAACCTTCA 1268
Db 384 oAlaAspLeuArgArgGlyGlyArgSerArgProAlaAlaArgGlyLeuGlnArgLeuHi 404
QY 1269 TTGTGGA-----CGCAAAACGCATCTTCAT-----CGGCT 1298
Db 404 sProGlyThrLeuArgArgTrpProArgGlnArgLeuHisHisProArgArgGlyArgGl 424
QY 1299 CATTCAACTCCACCCCGTTC---CGCAGGCTCAATACCGAATGGCGCTGCATCG 1355
Db 424 yAlaAlaProArgProGlyValProArgThrAlaAlaAlaAlaAlaAlaAlaAlaHisPr 444
QY 1356 AAAGCCCCAAATCGCAGAACAA-----GATGAGGCGCACCTCGCCGATACACAC 1406
Db 444 oLeuProProValGlyArgThrAlaAlaGlyArgSerAlaGlnProAlaAlaGlyArgTh 464
```

```
QY 1407 CCGAATACGCTACCGCTTACCTCGA-----CAAACACAACGCGCTCAATGCACG 1460
Db 464 rGlyAlaArgSerProGlyArgProArgGlyLeuAlaGlySerAlaProAlaMetSerAr 484
QY 1461 A-----TC 1463
Db 484 gArgGlnThrCysSerLeuLeuLeuAlaPheGlyLeuPheTyrLeuValProLeuSe 504
QY 1464 CGCGCACCGGAAAAAACCCTACCGAAGCAACGACCCGA 1497
Db 504 rAsnHisGlyLeuTrpIleProAspGluThrArg 515

RESULT 21
US-09-252-991A-20186
; Sequence 20186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20186
; LENGTH: 1098
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20186

Alignment Scores:
Pred. No.: 6,898-14 Length: 1098
Score: 280.00 Matches: 171
Percent Similarity: 35.7% Conservative: 43
Best Local Similarity: 28.5% Mismatches: 195
Query Match: 9.8% Indels: 190
DB: 2 Gaps: 39

US-10-665-990A-13 (1-1561) x US-09-252-991A-20186 (1-1098)
QY 136 CAATACTTCCAAACCTGCTCTCTGGACAAACATCTCGCAAAATCCGGCACACCCCTCATAA 195
Db 3 ArgTyrPheArgAlaProGluProPheSerProProArgValThrGluGluSerGln 22
QY 196 CACCGGCTATCCGA-----CATCTACCTGCTCGACGACCC----- 231
Db 23 ProArgLysIleHisTrpSerValProHisValProValThrArgProValAlaAlaGln 42
QY 232 -----CCAGAAAGCCTTGGCGCGCGCGC-----CGCCCTTATCGAATC 270
Db 43 ProArgProAlaProArgLeuProArgSerArgProGlyPheArgArgGlnArgLeu 62
QY 271 TGCCGAACACAGCCT---CGATT---GCAATATCATATTTCGCGCAACGACATTCCCG 324
Db 63 ArgArgArgAspProHisArgTyrProProValLeuProAlaAlaArgAlaGly 82
QY 325 CAGGCTGTGTTCACCTCATGTACCTTCCCGCAGACG----- 363
Db 83 AlaGlyAlaGluProProArgProProGlyArgArgAlaAspArgLeuGlnProGlyThr 102
QY 364 ---CGCGCTGCGCGTACGCTCTTGGACCAACACACGCGCGGTTCGACGATCT 420
Db 103 AspArgArgAlaAlaGlnProGlyAlaGlyArgArgValHisPro-----GlyGlySer 120
QY 421 CTGCTCGCCTCGACGCAATCCCAATATCGAATGCGCGCTGTTCACCCCTTCGT--- 477
Db 121 ProGlyArgProAlaArg-----ArgLeuArgProGlyGlyAlaArgProArgArg 138
```

[illegible]

```
Db      138 AlaArgGlyValasp-----lIleArg 144
QY      461 CTGTTCAACCCCTTCGTCTACGCAAAAGCGCGCACTCGGTACTGTACCGACTTCCCC 520
Db      145 AlaPheLysProPheSerLeuSer-----ArgGlyLeuArgHisLeuLeu----- 159
QY      521 CGCTCAACCGCGCATGCACAAACAAATCTTTACCGCGCAACACCGCGCCACCATATCTC 580
Db      160 -----ArgArgAspHisArgLysIleLeuValValAspGlyGluValAlaPheThr 176
QY      581 GCGGACGCAATATATGCGGACGAATACTTTCAAAGTCGGTAGGACACACCGTTTTTCGCCGAC 640
Db      177 GlyGlyValAsnIleSerAlaHisIleAlaProAlaGluMetGlyAlaIleAlaIlePheArgasp 196
QY      641 CTGGACATCTCTCCGACCGGAGCGTCTGGCGAGTAGTATCGACGACTTCGACCGCTAC 700
Db      197 AspValLeuArgIleGluGlyProAlaValHisGluLeuGluArgCysPheSerAlaThr 216
QY      701 TGGGCAGCCATTCGCCCCACAAACGCCACGCATCATCCGACGGGGGACACATCGGCAAG 760
Db      217 Trp-----ArgMetMetPheGlnGlyArgPhe----- 225
QY      761 GGTCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACGCGTCTCGCTACCGC 820
Db      226 -----PheArgLeuThrArgArgLeu 232
QY      821 GAAACCGTCGAACAGTCGCCCTCTTACCAAAAAATACAGCGGGACGCATCGACTGGCAG 880
Db      233 GluArgLeuArgAsnProPro----- 239
QY      881 AGGCTCAAAACCGGCTGATACGACGACCCCTGCAAAAGGA-----CTC 925
Db      240 -----ProArgArgGlyAlaValGlyLeuValVal 249
QY      926 GACCGGACCGCGCAAAACCGCGGATTCGCGGAGGCTGCAAGACGCGCTCAAAACAGCC 985
Db      250 LeuSerSerArgArgSer-----IleHisArgAlaTyrLeuHisAlaIleArgAla 267
QY      986 GAAAAAGCGTCTATCTGTTCACCCCTATTTGCTCCCTTACAAATTCGGGCACAGACGA 1045
Db      268 ArgArgSerValLeuValAlaAlaTyrPheIleProAspArgArgMetValMetAla 287
QY      1046 CTGGCAAAACTGGTCAGACGCGCATAGAGTTCACGCTCTGACCACTCGCTACAGGG 1105
Db      288 LeuArgGluAlaAlaArgArgGlyValGluValHisLeuLeuLeuAsnAla-----Arg 305
QY      1106 ACCGAGCTTTCGCGCGCTCCATTCGCGTACGTCAAATACCGAAACCGCTGTCAAAGCC 1165
Db      306 SerAspHisProIleLeuGluPheMetAlaArgAlaPheTyrGluArgLeuLeuGlyAla 325
QY      1166 GGCATCAAACTTACGAGTGCAAACCCAAACCATGCGCTCCCGCCACAAAGACAAAGGC 1225
Db      326 GlyValArgIlePheGluTrpGln-----ArgGly 335
QY      1226 CTGACCGGGAGTCCGTAACAGCTGCATGCGCAAAACCTTCATTTGTGACGGCAACGC 1285
Db      336 Val-----LeuHisAlaLysThrAlaValValAspGlyValTrp 348
QY      1286 ATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGGC 1345
Db      349 GlyThrIleGlySerPheAsnLeuGluArgLeuSerLeuAlaPheAsnHisGluValAsn 368
QY      1346 GTCGTCATCGAAAGCCCCAAATTCGCAACAGATGGAGCGGACCCCTCCCGCATCCACA 1405
Db      369 AlaValPheAlaAspProArgLeuGlyGlnGlnLeuGluAspSerPheArgGlyAspCys 388
QY      1406 CCGGATACCGCTACCGCTTACCTCGACAAACACACCGCTGCAATGGCACGATCCC 1465
Db      389 GlyAsp---CysArgGluValThrLeuAlaGluPheArgArg----- 401
QY      1466 GCCACCCGAAAAACCTTACCCGAAACGAAACCCGAAACCTTTGGAAACGCATCCGCGCA 1525
Db      402 -----ArgProLeuTrpGlnLysLeuLeuGlu 410
```

```
QY      1526 AAAATCCTATCCTCTGCTG 1543
Db      411 ArgAlaLeuSerLeuLeu 416

RESULT 23
US-09-976-594-615
; Sequence 615, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 615
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 2789525CD1
US-09-976-594-615

Alignment Scores:
Pred. No.:      1,76e-13      Length:      904
Score:          374.50      Matches:    133
Percent Similarity: 37.7%      Conservative: 60
Best Local Similarity: 26.0%      Mismatches:  204
Query Match:      9.6%      Indels:      115
DB:              2          Gaps:        23

US-10-665-990A-13 (1-1561) x US-09-976-594-615 (1-904)
QY      217 CTGTCGACGACCCCAACGAGCCCTTTCGCGCCGCGCCGCTTATCGAATCTGCCGA 276
Db      172 ProArgArgArgLysSerArgSerProArgArgArgSerProValArgArg 191
QY      277 ACACAGCCTCGATTGTCATCTACTACTATTGGCGCAACGACATTTCCGCGAGCTGCTGT 336
Db      192 GluArgLysArg-----SerHisSerArgSer----- 200
QY      337 CAACCTCATGTACCTTGGCGAGAACGCGG---CGTGGCGGTACGCTGCTGTGGACGA 393
Db      201 -----ProArgHisArgThrLysSerArgSerProSerProAlaProGluLys 216
QY      394 CAACAACACGCGCGGTTGGAGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGA 453
Db      217 LysGlu-----LysThrProGluLeuProGluProSerValLysValLys 231
QY      454 AGTGGCGCTGTTCAACCCCTTCGTCTCTAGC---CAAATGGCGCGCACTCGGTACTGTGAC 510
Db      232 GluProSerValGlnGluAlaThrSerThrSerAspIleLeuLysValProLysProGlu 251
QY      511 CGACTTCCCGCTCAACCGCGCATGCACAAACAAATCCTTTACCGCGCAACACCGCGC 570
Db      252 ProIleProGluProLysGluProSerProGluLysAsnSerLysLysGluLys 271
QY      571 CACCATACTCGG---CGAGCGCAATATCGCGCAGCAATACTTCAAAGTCGGTGAGGACAC 627
Db      272 GluLysThrArgProArgSerArgSerLys-----SerArgSerArgThr 288
QY      628 CGTTTTCCGCGACCTGGACATCTCGCCACCGGACGCT----- 666
Db      289 ArgSerArgSerProSerHisThrArgProArgArgHisArgSerArgSerArgSer 308
QY      667 -----CGTGGCGGAAGTATCGACGACTTCGACGCTACTGGGCAAGCCATTCGCG 717
```


Qy	643	-----GGACAT-----	648
Db	536	AlaAlaValGlyGlyAlaGlyLeuArgGlyHisArgProGlyArgTrpProValGly	555
Qy	649	-----CCTGCCACCGCAGCGTCTCGGCAAGTATCGCAGCACTT	690
Db	556	LeuValArgArgLeuProAlaGlyArgGlyThrArgGlnGlyArgThrArg	575
Qy	691	CGACCG	696
Db	576	ArgProGlyProGlyProAlaProGlyAlaGlyAlaAspArgProProAlaPhe	595
Qy	697	CTACTGGGCAAGCAATTCGCCCCACACGCG	729
Db	596	LeuLeuSerGlyAlaAlaArgProArgArgAlaGlyProAspArgGlnProProAla	615
Qy	730	GCGCATCATCGCAGCGCAACATCGG	756
Db	616	GlyProAspProGlyArgArgHisArgGlyAlaValThrArgCysGlyAlaGlyThrGlu	635
Qy	757	-----CAAGGTCCTCAAGCACTCGGATACACGA	786
Db	636	AlaGlyValProArgProProGlnProAlaAspGlyAlaThrAlaAlaAspGlnGlnGly	655
Qy	787	CGAAACATCATCAGACACCGCTCTCGCTACCGCAACCGCTCGAACAGTCGCCCTCTA	846
Db	656	ArgAspProGlySerHisProProGlyArgAspAlaAlaArgAlaThrValArgProArg	675
Qy	847	CCAAAAAATACAGACGGG	873
Db	676	ProSerArgProAlaGlyGlnTyHisProProAlaArgValAlaGlyLysArgHisArg	695
Qy	874	CTGCGACAGCGTCCAAACCCGCTGATCAGCAGCACCCCTGCCAAAGGACTCGACCGCGA	933
Db	696	LeuLeuAlaGluThrArgArgProGlyLeuArgArgProAlaArgAlaAspArgArgThr	715
Qy	934	CCGCGG	963
Db	716	ProAlaAspProAspLeuAspProLeuGlyThrLeuGlnLeuHisProValArgThrTyr	735
Qy	964	GCAAGACGCGTCAAAACGCGCGGAAAAAGGCTATCTGTTTCAACCTATTTCGTC	1023
Db	736	AlaGluArgGlnProGlyAlaGlyAlaAspArg	750
Qy	1024	TACAAAATCCGGCAGACAGCACTGGCAAAACTGGTCAGAGCAGCGCATAGACGTATTACCGT	1083
Db	751	-----AlaLeuArgHisGlnArgArgGlnGlnAlaGlyValGlyLeuHisArgGlnAlaThr	769
Qy	1084	CCTGACCAACTCGCTACAGGGACCGAGTTGCGCGCTGCATTCCGGCTACGTCAAATA	1143
Db	770	GlyAlaGlyGlnLeuProGlyAspArgAlaAspProProArgLeuTrpArgHisGlu	789
Qy	1144	CCGAAAACCGCTGCTCAAAGCGCGCATCAAACTCTACGA	1182
Db	790	ArgArgArgAlaAlaGlnGlyGlySerGlyAlaAlaArgPheHisProProLeuGlyHis	809
Qy	1183	-----GCTGCAACCCCAACATCGCTCCCGCCACAAAGACAAAGCGCT	1227
Db	810	ArgLeuProGlyArgGlnPro	828
Qy	1228	GACCGGAGCTCCGTAAACCGCTGCATGCCAAAACCTCATTTGTCGACGCGCAACCGCAT	1287
Db	829	GlyThrAlaLeuPheGlyGlnProLeuArgGlnLeuLeuHisArgArgAlaAlaHis	848
Qy	1288	CTTCATCGGCTCATTCAA	1329
Db	849	LeuGlnGlnLeuProGlnGlyArgGlnArgProProAlaAspAlaAlaGlySerProAla	868
Qy	1330	-----CAATACCGA	1338
Db	869	ArgIleAspGlnProAlaLeuArgAlaProAlaAlaArgProAspProProArgTyrLeu	888
Qy	1339	AATGGCGTCTGTCGAAAGCCCAAAATCCGACAGACAGATGGAGCGCACCTTCGCGCGA	1398

Db	889	ProGluArgGluGlnGlyAlaValAlaGlyArgGlnGlyProAlaProArg	--- 907
Qy	1399	TACCACACCCGAATACGCTTACCGCGT---	TACCCTCGACAAACACAACCG----- 1446
Db	908	-----ArgLeuProArgProPheArgGlnGlyLysProGlyLeuPro	922
Qy	1447	-----CCTGCATGGCAGCATCCGCCACCCGAAACCTTACCC	1485
Db	923	AlaProLeuLeuGlyGlnValProArgGlnGlyArgGlnProAlaProGlyAspLeuPro	942
Qy	1486	GAACGAACCCGAGCCAAACTTTGGAAACGATCGCGCGCAAAAAATCCTATCCCTGTGTGCC	1545
Db	943	ArgArgProAlaSer---LeuAlaGlyAlaAlaGlyArgAspProProLeuProAlaAla	961
Qy	1546	CATCGAAGG	1554
Db	962	ThrGlyArg	964
RESULT 25			
US-09-252-991A-25018			
; Sequence 25018, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 25018			
; LENGTH: 1225			
; TYPE: PRT			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-25018			
Alignment Scores:			
Pred. No.: 3,13e-13 Length: 1225			
Score: 272.00 Matches: 176			
Percent Similarity: 34.2% Conservative: 46			
Best Local Similarity: 27.1% Mismatches: 179			
Query Match: 9.5% Indels: 248			
DB: 2 Gaps: 39			
US-10-665-990A-13 (1-1561) x US-09-252-991A-25018 (1-1225)			
Qy	157	CCTGGACACAATCTCGCAATTCGGCA-----CACCCCTCATAACAA	198
Db	172	ProAlaArgProGlyAlaProAlaAlaLeuAlaValAlaProGlyProGlyLeuGln	191
Qy	199	CGG-----GCTATCCGACATCTACCTCTCGAGACCCCCACGAAGCCCT-----	243
Db	192	ArgPheGlyAlaAspGlnArgAlaValAlaAlaArgProGluSerGlnProArgAlaGly	211
Qy	244	-----TGCCGCCCGCGCGCCGCT 261	
Db	212	AlaGlyGluArgGlnAlaProSerArgAspArgGlnHisGlnCysGlnArgAlaAla	231
Qy	262	TATCGAATCTCCGAAACACAGCCTCGATTGTCGAATCTACATTGGCGCAACGACATTTC	321
Db	232	GluArgLeuHisArgGlyArgProGlyTyr-----HisProAlaAlaArgArgHis---	247
Qy	322	CGGAGGCTGCTGTTCAACCTCATGTACCTTGCCGCGAAGACGCGCGCTGCGCGTACGCT	381
Db	248	-----ArgAlaHis-----ArgGlyAlaAlaArgArgArgLeuGly	259
Qy	382	GCTGTT-----GGAGACAACAACACGCGCGGTTGGACGATCT	420

Db 260 AlaValArgLeuGlyArgHisArgGlyAspGlnArgHisProGlnValGly----- 277
QY 421 CTGCTCCCTCGACAGCCATCCAATATCGAAGTCGCCTGTTCAACCCCTTCGTCT 480
Db 278 -----ArgProArgArgProAlaArg-----LeuArgGln 287
QY 481 ACGCAATGGCGCGACCTCGCTACCTGAC----- 510
Db 288 ArgSerLeuArgArgTArgProAspProGluArgArgGlyGluTrpArgProAla 307
QY 511 -----CGACTTCCCGCCCTCAA-----CCG 531
Db 308 PheArgArgGluArgPheProGlnProGlnArgAlaValProArgThrGlyAlaAspHis 327
QY 532 CCGCATGCACAAACATCTTTACCGC-----CGACAACCGCGCCAC----- 573
Db 328 ProArgArgHisArgProAlaTyArgAlaLeuArgGlnProValHisArgArgProPhe 347
QY 574 -----CATACTCGCGGACGCGCAATATCGCGACGAATA 606
Db 348 AlaAlaProProGlyAlaGlyLeuGlnArgArgGlyAlaGlyArgGlnArgAla 367
QY 607 CTTCAAAGTCGTCGAGACACCGTTTCGCCGA----- 639
Db 368 ValGlnLeuArg-----HisLeuTyProSerLeuGlyValLeuGlyAlaAspLeuPro 385
QY 640 -----CCTGGA-----CATCTCGCGCACCGGACGCT 666
Db 386 ValAlaValProGlyAlaAlaThrLeuSerGluArgLeuHisProAlaHisGlnArg 405
QY 667 CGT-----CGCGAAGTATCGCAGCATTCGA 693
Db 406 ArgGlyArgLeuPheAlaAspProArgArgProTrpArgArgAlaValArgLeuGly 425
QY 694 CCGCTACTGGCAGCATTC-----CGCCACACCCCGCGCATCATCCG 741
Db 426 -----LeuGlyProGlnPheAspLeuArgCysArgProGlyAspArgHisGlyPro 443
QY 742 CAGCGCAACATCGGCAAGGTCTTCAAGCACTCGGATACACGACGAAACATCCAGACA 801
Db 444 PheAspGlnProArg----- 448
QY 802 CGCGCTCTGCTACCGCGAACCCTCGACAGTCGCCCTCTTCACCAAAAAATACAGAC 861
Db 449 -----ProLeuArgArgAspArgArgPheAlaAlaLeuLeuArgProGly----- 463
QY 862 GGGACGCATCGACTGGCAGACGCTCCAAACCGCCTGATCAGCA-----CACCCCTCGCAA 918
Db 464 -----AlaLeuGlnGlyHisProValAspGlnGlnSerArgProAlaPro 478
QY 919 AGGACTCGACCGGACCGCGCAACCCCGATTGCGGGAG----- 960
Db 479 ArgLeuArgProArg-----LeuAlaAlaGlyAlaAlaGluPheLeuLeuGlyProGly 496
QY 961 GCTGCAAGACGCGCTCAACAGCCGCAAAAGCGTCTATCTGTTTCACCCCTATTTCGT 1020
Db 497 AlaAlaProArgAlaLeuArgGlyArgGlyArg-----SerLeuValLeuLeuGlnArg 515
QY 1021 CCCTACAAATCCGACACAGCAGCTGGCAAACTGTCGAGACGCGCATACGCT--- 1077
Db 516 -----ArgPheGlnGlyThrProArgAlaGlyAlaGly-----HisArgArgThr 530
QY 1078 -----TACGTCCTGACCAA-----CTC 1095
Db 531 ValGluProArgArgAlaArgArgLeuSerArgProValHisArgProArgArgAlaLeu 550
QY 1096 GTTACAGCGGAC-----CGA 1110
Db 551 AlaThrGlyAspGlyAlaLeuArgAlaLeuGlnArgLeuArgGlnHisGlnArg 570
QY 1111 GTTTCGCGCGTTCATTCGCGTACGTCAATACCGAAACCGCTGCTCAAGCGCGCAT 1170
Db 571 GlnAlaArgGlnProLeuGlnProArgGlyIleGlyAlaGlyGlnSerGlnGln 590

QY 1171 CAACATCTACGAGCTGCAACC-----CAACCATGCCGTCCCGCCAC 1212
Db 591 ArgLeuProArgAlaValAlaGlyProGlyGlnLeuHisLysProGlyArgLeuAlaAsp 610
QY 1213 AAAAGACAAAGCGCTGAC----- 1230
Db 611 HisArgGlnArgProAlaGlyGlyGlnLeuAlaGlyGlyGlnAlaAlaArgArgArgArg 630
QY 1231 -----CGCAGCTCCGTAAACAGCTGCATGCCAAAACCTT----- 1266
Db 631 ProGlyAlaGlyLysIleAspGlnLeuGlnProArgProGlyAlaArgSerLeuAlaGly 650
QY 1267 ---CATTGTGGAGCGCAACGCTCTTCATCGCTCATTCACACCTCGACCCCG----- 1317
Db 651 ProGluProGlyHisArgArgLeuSerHisArg-----AspProArgProHisArgArg 668
QY 1318 -----TTCCGACGCTCAATACCGCAATGGCGTCTCATCGAAAGCCCCCAA 1365
Db 669 ArgCysHisLeuPheArgThrAlaGlyAspArgCysAlaAlaArgArgAspGlnHis 688
QY 1366 AATCGCAGACAGAT---GGAGCGCACCTCGCGGATACACACCCGAATACGCTTACCG 1422
Db 689 ProGlyArgThrAspProGlyGlnHisPro-----LeuHisAspGlnArgArgHis 706
QY 1423 CGTTACCTCTGA-----CAACACACCGCTCAATGGCAGCATCCCGCCACCCG 1473
Db 707 ProTyProArgProArgProHisArgGlnLeuProAspProAlaArgValGlyPro 726
QY 1474 AAAAACCCTACCCGAAACCGAACCAACCTTTGGAAACGCGATCCCGCAAAAATCCT 1533
Db 727 AspArgLeuAlaProGlyArgGlnArgGlnSerHisGluAlaGlyGlnGluSerSerArg 746
QY 1534 ATC-----CCTGCTGCCATCGAAG 1554
Db 747 AlaGlnArgProAlaAlaGlnArg 755

RESULT 26
US-09-252-991A-26099
; Sequence 26099, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26099
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26099

Alignment Scores:
Pred. No.: 3,95e-13 Length: 863
Score: 270.00 Matches: 174
Percent Similarity: 32.0% Conservative: 36
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 9.5% Indels: 242
DB: 2 Gaps: 33

US-10-665-990A-13 (1-1561) x US-09-252-991A-26099 (1-863)

QY 208 CGACATCTACCTCTCGACGACGCCCGCCCGCGCGCCCTTATCGA 267
Db 45 ArgArgArgProSerGlnArgArgGlyAspProArgLeuArgLeuProLeuArg 64

```
QY 268 ATCTCCGAACACAGCCT-----CGATTGCAATACTACAT----- 303
Db 65 GluArgArgGileArgProCysLeuArgSerArgArgAspCysPheHisArgProGlyCys 84
QY 304 -----TTGGCGCAACGACATTTC-----CGGCAG 327
Db 85 GlySerHisProProAspGlyGlnAlaProValGluAspArgHisAlaArgArg 104
QY 328 GCTGCTGTTCAACCTCATGCTCTGCGCGCAGAACCGCGCTGCGGTACGCTCTGCTTT 387
Db 105 GlyAlaValHisProArgLeuArgArgArgThrGlyArg-----ArgAspProAlaAla 123
QY 388 -----CGACGACAA 396
Db 124 ArgGlyArgThrHisArgLeuProThrAspAspGlnGlyAlnArgArgArgPro 143
QY 397 CAACACGCG----- 405
Db 144 ArgHisAlaProGlyGlyGlyThrArgProThrAlaGlyProValAlaHisArgThrLeu 163
QY 406 -----CGGTTGGACGATCTCTCTGCTCGCTCGACAGCCATCCCAATATCGAAGTGGCT 462
Db 164 GlySerAlaGlyArgLeuArgGlnArgArgThrAspProGlyGluGlyArgAlaAlaPro 183
QY 463 GTT-----CAACCTTCGTCCTACGCAATGGCGCG 495
Db 184 AlaProCysGlyAspProGlyValArgArgProProArgGlnProAlaLeuProArgArg 203
QY 496 ACTCGGCTACCTGACCGACTTCCCGCTCAACCGCGCATCGCAACAACATC----- 549
Db 204 ThrArgLeuLeuGlyAlaAlaProProProGluGlyHisArgGlySerSerLeuProGly 223
QY 550 -----CTTTACCGCGCAACCGCGCACCATATCTCGCGCGACGCAATATCGCGCAGCAATA 606
Db 224 AlaArgCysArgProAlaSerArgHis----- 232
QY 607 CTTCAAGTCGTGGAGGACACCGTTTTCGCGACCTGGACATCTCGCCACCGGAGCGT 666
Db 233 -----GlyArgGlyGlyGlyGluGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 250
QY 667 CGTGGCGGAAGTATCGCAGCTTCGACCGCTACTG-----GGCAAGCGCATTCGCGCCA 720
Db 251 ArgArgValSerSerArgArgGlyTrpProLeuLeuLeuProGlyAspGluTrpProPro 270
QY 721 -----CAACGCCACGCGCATATCCG-----CAGCGG 747
Db 271 AlaGlyArgThrProGlyHisArgThrGlyHisArgProGlyProGlyGlyLeuAlaAsp 290
QY 748 CAACATCGGCAAGGTCTTCAAGCACTCGGATACAA----- 783
Db 291 ProArgArgGlyProSerAlaAlaAlaAlaAlaGlyThrGlyGlyThrAlaArgAla 310
QY 784 -----CGACGAAACATCCAGACGCGCTCTCGCTACCGCA 822
Db 311 CysHisArgGlyProProLeuArgArgGlyArgArgSerArgLeuProAlaAlaAspArg 330
QY 823 AAC-----CGTCGAACAGTCGCCCTCTACCAAAA 852
Db 331 HisGlyAlaAlaLeuGlyProGlyThrAlaAlaArgArgThrHisArgProTrpProGly 350
QY 853 AATACAGACGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCAGCACCCC 912
Db 351 ArgGlyAlaGlyAspAlaLeuLeu-----ArgProAsp-----AlaGlyGlnAlaAspArg 368
QY 913 TGCAAAAGGACTCGACCGCGACCGCGCAACCGGATTGCGGGAGGCTCGAAGACGC 972
Db 369 LeuArgGlyAspProArgArgSerProAlaGlnAlaAspProArgAlaGlyAlaLeuArg 388
QY 973 GCTCAAAACGCCGAAAAAG-----CGTCTATCTGTTTC 1008
Db 389 AlaAlaGlyGlyArgArgGlnProAlaValProArgGlnProProGlyAlaSerGlyLeu 408
QY 1009 ACCCTATTTTCCTCCCTACAAAATCCGGCAGACGCACT----- 1047
```

```
Db 409 ArgArgArgArgGlyAspHisArgLeuHisArgThrLeuArgArgArgProGlnPro 428
QY 1048 -----GCCAAAACCTGGTCAGCAGCGCATAGACGT----- 1077
Db 429 AlaThrAlaProAlaGlyArgArgGlyThrGlyProGlyArgArgProAlaLeuProGly 448
QY 1078 -----TACGCTCTGACCAACTC----- 1095
Db 449 ArgArgArgGlyGlnCysAlaProAlaArgProGlyArgLeuAlaGlnArgArgArgGly 468
QY 1096 -----GTCACAGCGCACCGA----- 1110
Db 469 ThrValGluAspArgProGluAlaArgArgThrAlaArgGlyAspArgProGlyArgPro 488
QY 1111 -----CGTTGCGCGCT----- 1122
Db 489 ArgLysArgSerProAlaAlaThrAlaGlyAlaProGlyAlaThrArgAspArgAla 508
QY 1123 CCATTCCGGCTACGTCAAAATACCGNAACCGCTGCTCAAGCCGCATCAAACTCTAGCA 1182
Db 509 ProAlaArgLeuArg-----ArgThrLeuGlyGlnArg-----ThrArgArg 523
QY 1183 GCTGCAACCCCAACCATGCGTCCCGCCCAAAAAGACAA-----AGGCTCTGAC 1230
Db 524 HisSerProProProArgValProArgGlyArgArgAlaProLeuGlyLeuArgProArg 543
QY 1231 CGCGAGCTCCGTACACCGCTGCATGCGCAAAACCTTCATTGTGGACGGCAACCGATCTT 1290
Db 544 ArgGlnProArgThr-----GlyArg-----GlyHisPro 553
QY 1291 CATCGGCTCATTTCAACCT-----CGACCCCGTTCGCGACGGCTCAATCCGAAATGGCGT 1347
Db 554 ArgArgArgGlnArgProGlyArgArgGlnPheArgHisAlaGlnGlyAlaAspGlyArg 573
QY 1348 GCTCATCGAAAGCCCAAAATCGCAGA----- 1374
Db 574 ArgHisArgArgGlyAlaGlyArgArgGlyArgGlnGlyArgGlnGlyProThrAlaAla 593
QY 1375 -----ACAGATGGAGCGCACCTCCCGCATACCACCGCAATACGCCTA 1419
Db 594 GlyAlaGlyGlyAsnGluAspGlyThr---ProAlaGluGlyArgSerArgArgGly 612
QY 1420 CGCGCTTACCTCCGACAAACACACCG-----CCTGCAATGCA--- 1458
Db 613 Pro-----ProArgProGlyArgProTrpArgAlaGlyAlaGlnProAlaGlyAlaGly 630
QY 1459 -----CGATCCCGCCACCCGAAAAACCTACCGAAGCAACCCGAAAGCCAA 1503
Db 631 ArgGlyGlySerArgArgLeuArgHisProAlaAlaHisProGlnArgGlnArgGlnGlu 650
QY 1504 ACTTTGGAAACGATCGCGCGCAAAATCTATCCCTGCTGCGCCATCGAAGG 1554
Db 651 AsnGlyAspAlaHisValThrAlaArgGlnAspProValHisHisArgArg 667

RESULT 27
US-09-252-991A-18427
; Sequence 18427, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18427
; LENGTH: 748
```



```

; SEQ ID NO 19467
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19467

Alignment Scores:
Pred. No.:      3,76e-13      Length:      470
Score:          269.00      Matches:    140
Percent Similarity: 33.6%      Conservative: 33
Best Local Similarity: 27.2%      Mismatches: 192
Query Match:      9.4%      Indels:    150
DB:              2      Gaps:      25

US-10-665-990A-13 (1-1561) x US-09-252-991A-19467 (1-470)

Qy 76 TCTCTGTTTCATGTTCTTCATGTTTCGCCCTTCCGACGAGCGGAAACCGTCATTT 135
    |||||::: ||| ::||| ||||
Db 15 SexLeuTyRGlyPheGlnSerIleAlaLaProGlyAlaLeu----- 280
Qy 136 CAATACTTCCAAACGTCTCTCTCGACAAACATCTCGAAATCCGGCACACCTCATAA 195
    ::: |||||::: ||||
Db 29 -----SerArgAlaProGlyArgGlyAlaGlyArgGlyIleHis----- 41
Qy 196 CAACGGGCTATCGACATCTACTCTGCTCGACAGCCCCACGAAAGCCTTCGCCGCCGCG 255
    |||||::: ||||| ||||| |||||
Db 42 -----ArgArgAlaGlyArgProGlyArgProAla 52
Qy 256 CGCCCTTATCGAATCTGCCGACACAGCCTCGATTGCAATCTACATTTGGCGCAACGA 315
    |||||::: ||||| ||||| |||||
Db 53 AspProAlaArgLeuCysArgProGlnPro-----AlaThrAspArg 66
Qy 316 CATTTCCGGCAGGCTGCTGTTCAACCTCATGTACCTTGGCGCAGAACGCG----- 366
    ||||| ||||| ||||| |||||
Db 67 PropHeHisGlnLeu-----TyArgArgSerArgGlyGlyLeu 79
Qy 367 -----CGTGGGTAGCGCTGCTGTGGACACAAACACGCGCGGTT 411
    ||||| ||||| ||||| |||||
Db 80 ArgArgCysProGlyArgArgGlyValProGlyAlaAspArgLeuGlnArgGlyLeu 99
Qy 412 GGACGATCTCTGCTCGCCT-----CGACAGCATCCCAATATCGAATGCGCCTGTT 465
    ||||| ||||| ||||| |||||
Db 100 GlnArgHisArgLeuArgProGlyThrAspProGlyGlnLeuAlaSerGlyLysHis 119
Qy 466 CAACCCCTTCGTCCTACGCAAAATGCG-----CGC 495
    ||||| ||||| ||||| |||||
Db 120 ArgAlaAlaAlaProAlaGlyThrAlaThrGlyAlaLeuAlaAlaGlnArgAlaAspArg 139
Qy 496 ACTCGGTACTACCGACTTCCCGCGCTCAACCGCGCATGCAACAATCTCTTAC 555
    ||||| ||||| ||||| |||||
Db 140 ThrThrAspProAlaArgProGlnAspArgPheGlyGlyGlyGlyLeu----- 157
Qy 556 CGCGGACACCG-----CGCCACCATCTATCGGCGAGC 588
    ||||| ||||| ||||| |||||
Db 158 ArgArgGlnProAlaAlaLeuProAlaHisArgPheArgArgHisGlyProValProAla 177
Qy 589 CAATATCG-----CGACGAATACTTCAAAGTCGTGAGGACACCGTTTTCCGCCACCT 642
    ||||| ||||| ||||| |||||
Db 178 AlaTyArgProGlyArgLeuLeuProValAlaAlaAlaGlyThrArgTyArgArgPro 197
Qy 643 GGACATCTCTCGCCACCGGAGCGTCTCGGCGAAGTATCGCACGACTTCGACCGCTACTG 702
    ||||| ||||| ||||| |||||
Db 198 AlaArgProThrGlyGlu----- 203
Qy 703 GGCAGCGCATTCGCCCAACACGCCACGGGCAATCATCCGAGCGGCAACATTCGGCAAGG 762
    |||||::: ||||| ||||| |||||
Db 204 GlyArgAlaTyArgAlaProArgThrGlyPhe-----ArgGluAlaArgThrGly 220
Qy 763 TCTTCAAGCACTCGGATACACGACGAAACATCCAGACCGCGTCTCTCGGCTACCGCGA 822
    ||||| ||||| ||||| |||||
Db 221 SecArgHisArgArgAlaGluProThrGlyArgGlyLeuGlnArgProAla-----Arg 238
Qy 823 AACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGACGGGACGATCGACTGGCAGAG 882

```

Best Local Similarity:	29.0%	Mismatches:	166
Query Match:	9.4%	Indels:	169
DB:	2	Gaps:	35

US-10-665-990A-13 (1-1561) x US-09-252-991A-24941 (1-475)

145	QY	CAAACTGTGCT-----CCTGGACACAATCCTCGCAAT	177
146			
147	Db	GlnProCysProGluArgLeuProTyrLeuGlyArgProArgGlnGlyProGlyCys	47
148			
149	QY	CCGGCACACCCCTCATAACACAGGGCTATCCGACATCTA-----CCTGCT	222
150		: :	
151	Db	ProArgGlnProGlyAspArgGleuValArgArgAlaGlyGlyProGlyAlaGly	67
152			
153	QY	CGACGACCCCACGAAGCCCTTGCCGCCGCCGCCCT---TATCGAATTCGCCGAACA	279
154		:	
155	Db	ArgArgValProGlnGlySerValProProGlyHisProGlyLeuArgArgGlnArgAla	87
156			
157	QY	CAGCCTCGATTGCAATACTACTATTTGGCGCAACGACATTTCCGGCAGGCTGTGTTCAA	339
158		: : :	
159	Db	GluProArg-----ArgArgSerArgArgHisGlyGlnArgAla---Gln	102
160			
161	QY	CCTCATGTACCTTGC CGAGAACC GCGCGTGC GTACGCGTGTGTGGACGACAA	399
162		: :	
163	Db	ProGluGlyAlaValArgGThrGlnArgAlaGlyAlaGlyGlnProGlyThrGlnArg	122
164			
165	QY	CAGCGCGGTTGGACGATCTCTGTCGCGCTCGACAGCCATCCCCAATATCGAAGTGC	459
166			
167	Db	-----GlyAspPro-----ProArgGln-----	128
168			
169	QY	CCTGTTCAACCCCTTCTGTCTACGCAATGGCGCGCACTCGGTACTCGTACCACGCTCCC	519
170			
171	Db	--ValGlnArgLeuArgGlnHisGlnGlnProGluIlePheValPro-----LeuPro	145
172			
173	QY	CCGCTCAACCGCGCATGCAACAACAAATCTTTACCGCGCAACACG-----	567
174		:	
175	Db	AlaValProValGlyProAlaArgArgLeuGlnArgGlnArgPheProCysAlaValAla	165
176			
177	QY	-----CGCCACATACTCGCGGAGCGCAATATCGGCGACGAATA	606
178		:	
179	Db	ValArgThrValGlnProAspLeuHisLeuTyrGlnArgGlnLeuArgArgProAla	185
180			
181	QY	CTTCAAAAGTCGTGAGGACCGTTTTTCGCGCACTCGACATCCT---CGCCACCGCGAG	663
182			
183	Db	ProValArg-----ArgArgProAlaGluProGlyArgHisArgGln	199
184			
185	QY	CGTCGTGGCGAAGTATCGACAGCTTCGACCGCTACTGGCGCAAGCCATTTCCGCGCCAAA	723
186		:	
187	Db	-----ProArgLeuArgProAlaVal-----LeuGln	208
188			
189	QY	CGCCACGCGCATATCGCAGCGGCAACATCGCAAGGGTCTTCAAGCATCTCGATACAA	783
190		: : :	
191	Db	ArgHis-----ArgArgGlnTyrArgPro--AlaThrGlyAsnArgAlaGln	223
192			
193	QY	CGACGAAACATCAGACACGCGTCTCTGCGCTACCGCGAAACCGTCAACAGATCGCGCCCT	843
194			
195	Db	-----TyrProGlyProGlyLeuPro-----AlaAlaAla	234
196			
197	QY	CTACCAAAAAATACAGACGGGACGATCGATCGCAGAGCGTCAAAACCGCCTGATCAG	903
198			
199	Db	ArgProPheArgArgProGlyLeu-----LeuValAspGlnAspArgGlnProAspArg	252
200			
201	QY	CGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATTCCCGGAGGCT	963
202			
203	Db	-----ArgValSerArgSerGlyAspLeuArgArgPro	263
204			
205	QY	GCA-----AGACGCGCTCAAAACAGCCGAAAAAGCGTCTATCTGTTGTTTCA	1017
206		: :	
207	Db	AlaGlyLeuArgArgThrHisArgAlaGlnGlyArgArgLeu-----	277
208			
209	QY	CGTCCCTACAAATATCGGCACACAGCCTTGGCAAAATCTGGTCAGACCGGCATAGACGT	1077
210			
211	Db	TyrArgSerArgArgHisArgThrGlyGlnProArgGlnGlyAspGlnArg	295
212			

	Qy	1078	TACGCTCCTGCACCAA-----CTCGTACA 1101 :::
	Db	296	ArgArgProGluProArgLeuSerPheProGlyGlnProLeuArgAlaValArgAlaArg 315 :
	Qy	1102	GCGCACGGAGTTGCCGCGTCATTCCGGCTA-----CGTCAATACCG 1146 :
	Db	316	ProAlaArgHisLeuArgValProLeuArgLeuProAlaAlaAspArgArgProValPro 335 :
	Qy	1147	AAAAACCGTGCTCAAAGCCGGCATCAAACGTCTACGAGCTCCAACCCAACCATCCGCTCCC 1206 :
	Db	336	-----GlyGlnArgArgArgLeuProGlyArgArgAspArgProLeuGluAla 352 :
	Qy	1207	GCGCACAAAAGA CA A AG CG CT G AC C GG GA GT CC GT T AAC C AG C CT GC AT GCC AAA ACC TT T 1266 ::
	Db	353	ArgArgLInArGHISleUGlUPROARGLLEUalagLYHISPROglu---GlnProLeu 371 ::
	Qy	1267	CATTGTGGACGGAACGGCATCTTCATCGGCTCATTTCAACCTCGACCCCCGGTTCGCGACG 1326
	Db	372	HIS-----GlnArgLeuGlnArgLeu---ArgProargGlnPro----- 383
	Qy	1327	GCTCAATACCGAAATGGCGGT-----CGTCATCGAAAGCCCC 1362
	Db	384	-----ArgGlnGlyArgPheValGluProLeuGlyProGlyArgGlnLeuProPro 400
	Qy	1363	CAAAATCGCAGAACAGATGAGGCGACACTCGCCGATACCAACCCGGAATACGCTTACCG 1422
	Db	401	GlnProArgala-----GlyAlaAspProArgGlyGluGlu-----Pro 413
	Qy	1423	CGTTACCTCTGA---CAAACAGACCGCTGCAATGGCAGCATCCGCCCAACCGAAAAAC 1479
	Db	414	ValArgProArgThrAlaValGlnGlnPrOaspLeu-:::
	Qy	1480	CTACCCGAACGAACCCGAAGCAAACTTTGGAAAACGCATCGCGCAAAAATCCTATCCCT 1539
	Db	427	LeuProGluaRgLeuarg-----Pro 433
	Qy	1540	GTGCGCCATCGA 1551
	Db	434	AlaLeuHiSArg 437

RESULT 30

```

US-09-252-991A-24311
; Sequence 24311, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIORITY APPLICATION NUMBER: US 60/074,788
; PRIORITY FILING DATE: 1998-02-18
; PRIORITY APPLICATION NUMBER: US 60/094,190
; PRIORITY FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24311
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24311

```

Alignment Scores:		
Pred. No.:	4, 48e-13	584
Score:	268.50	107
Percent Similarity:	37.6%	59
Best Local Similarity:	24.2%	167
Query Match:	9.4%	109
DB:	2	14
		Gaps: 1
		Indels: 0
		Mismatches: 1
		Conservative: 1
		Matches: 1
		Length: 1

US-10-665-990A-13 (1-1561) x US-09-252-991A-24311 (1-584)


```
QY 769 AGCACTCGGATACACGACGA-----AACATCCAGACAGCGCCTCTCGG 813
Db 403 ArgArgArgLeuGlnArgAlaAspAlaThrGlyAspSerLeuArgHisGlnArgProAla 422
QY 814 -----CTACCGGGAACCGTCGACAGTC-----GCCCT 843
Db 423 GluHisArgArgHisProArgGlnArgProThrValGlnArgAlaGlyAspLeuGluPro 442
QY 844 CTACCAAAAAATACAGACGGGACGCATCGACTGGCAGAGCGT----- 885
Db 443 LeuHisAlaGlu---ProGlyArgProArgLeuGluProArgLeuTrpArgArgGln 461
QY 886 CCAAAACCGG-----CCTGATCAGCAGACACCCCTGCAAAAGACTCGACGGACGG 936
Db 462 ProLeuProAlaGlyArgProAspLeuAlaHisAlaGluArgAlaThrAlaGlyArgPro 481
QY 937 CGCGAAACCGCGATTCGCGGAGGCTGCAAGACGCGCTCAACAGCCCGGAAAAAGCGT 996
Db 482 Leu-----TrpArgAspArgGlnLeuHisAla---LeuGlyArgSerGlnArg 496
QY 997 CTATCTGGTTTCACCCCTAFTTCGTCCTACAAA-----ATCCGGCAGACGCG 1044
Db 497 LeuAlaGlyMetAspGlyGlnArgArgValArgGlnGlnProAspGlnArgArgLeuArg 516
QY 1045 ACTGCGAAA-----ACTGTGCA-----GGA 1065
Db 517 ProGlyGlnHisGlnGlyLeuSerThrGlyAlaAspProLeuArgGluProThrAspGly 536
QY 1066 CGGCATAGAGTTTACGCTCCTGACCACTCGCTACAGGCGACCGGTTGCGCGCTCCA 1125
Db 537 GlnHisArgArgGlnArgProProAlaGlyAlaLeuGly-----ArgGly 551
QY 1126 TTCCGGCTACGTCAAATACGGAACCGCTGCT-----CAAAGC 1164
Db 552 LeuLeuProGlyGluValProAspArgAlaAlaGlyProSerGlyGlnArgGlnArgAla 571
QY 1165 CGGCATCAACTCTAGAGCTCAACC-----CAA 1194
Db 572 ArgSerArgAlaAlaArgGlyGlyThrProGlyGlnArgProAlaAlaGlyPheProAsp 591
QY 1195 CCATCGCGTCCCGCCACAAAAGACAA----- 1221
Db 592 ProCysArgSerArgArgGlnHisGlnProGlyGlyArgAlaArgThrAlaAla 611
QY 1222 ---AGCCTGACCGGAGCTCGCTAACCAAGCCTGATGCCAAACCTTCATTGTGACGG 1278
Db 612 GlyGlnProGlyArgArgAsnArgGlnArgProAlaAlaArgGlnArg---ArgLeuGlyArg 630
QY 1279 CAAAGCATCTTCATCGGCTCATCACTCACTCGACCCCGCTTCGACAGCGCTCAATACCGA 1338
Db 631 ProGlyLeuPheArgArgfileAlaLysArgGlnProAlaThrArgGlyProProArgArg 650
QY 1339 AATGGG-----CGTCGTCATCGAAAGCCCAAAAT 1368
Db 651 ProGlyLeuProGlyThrLeuProProGlyHisAlaGlnThrHisArgGlnPro---Gly 669
QY 1369 CGCAGAACAGATGGACGACCCCT---CGCCGATACCAACACCCGGAATACGCTACCGCGT 1425
Db 670 ArgProAlaAspLeuLeuGlyProAspArgArgTyr----- 681
QY 1426 TACCCTCGCAACACACACCGGCTGCAATGGCAGCATCCGCGCAC 1470
Db 682 -----AlaMetSerArgHis 686

RESULT 33
US-09-252-991A-27502
; Sequence 27502, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
```

```
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27502
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27502

Alignment Scores:
Pred. No.: 4,83e-13 Length: 441
Score: 267.50 Matches: 140
Percent Similarity: 36.2% Conservative: 48
Best Local Similarity: 26.9% Mismatches: 173
Query Match: 9.4% Indels: 159
DB: 2 Gaps: 28

US-10-665-990A-13 (1-1561) x US-09-252-991A-27502 (1-441)
QY 208 CGACATCTACCTGCTCGAGACCC---CCACGAAGCCCTTCCGCCCGCGCCCTTAT 264
Db 9 ArgAsnArgProAlaAlaArgValProLeuProGlyThrProValProGlnGlyProGly 28
QY 265 CGAATCTGCCGAACACAGCCTCGATTTCGAATACTACATTTTGGCGCAACGACATTTCCGG 324
Db 29 -----LeuArgArgThrAlaArg 34
QY 325 CAGGCTGCTGT---CAACTCATGTACTTCCCGCAGAACG----- 363
Db 35 GlnArgAlaAlaProGluProArgProGlySerGlyAlaArgThrLeuGlyLeuGly 54
QY 364 -----CGCGTGGCGGTACGCTGCTGTT 387
Db 55 ProGlyAspArgProAspValArgLeuAlaGlyLeuArgArgAlaAspArgProArgSer 74
QY 388 GGA-----CGACAAACACAGCGGGTTGGACCATCTCTGCTCGCCCTCGACAGCCA 441
Db 75 GlyGlyThrArgArgProHisArgAspLeuGluAlaProAlaProArgValArg 94
QY 442 TCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCCTCTACGAAATGGCGGCACTCGG 501
Db 95 HisArgAlaGlyGlyProProGlnProAlaGlyProAspArgGlnSerArgProHis 114
QY 502 CTACCTGAC-----CGACTTCCC-----CCG 522
Db 115 LeuProProAlaProGlnGlyArgArgLeuProLeuAspProLeuAlaArgProGlyPro 134
QY 523 CTTCAACCGCGCATGCACAAATCTTTTACCCCGCACCAACCGCGCCACCATCTCGG 582
Db 135 ProArgArgProGlyGlnAlaAlaLeuHisTrpArgGlyProArgHisHisProAla 154
QY 583 CGAGCGCAATATCGG-----CGAGCAATACTTCAAAGT 615
Db 155 ThrAlaGlnGlyArgProProAlaProGlyArgGlyGlyIleArgGlnHisProArgGly 174
QY 616 CGGTGAGGACACCGTTTTCGCGCAGCTGGACATCTCGCCACCGCGCGCTCGTGGCGA 675
Db 175 ArgSerGlyHisArgCys-----ProGlyGlyAspArgAlaArgGlnProPheLeuArg 192
QY 676 AGTATCGCAGCACTTCGACCGCTACTGGGCAAG-----CCA 711
Db 193 AlaHisHisArgLeu---ProLeuArgArgArgAlaArgGlnAspProGlyAspProPro 211
QY 712 TTCCGCCCAACACGCGCAGCGCATCATCG-----CAGCGCAACATCGGAAGG 762
Db 212 LeuArgProProGlyProGlyValLeuProAlaProValAlaGlyProAlaAlaGly 231
QY 763 TCTTCAAGCACTCGGATCAACGACGAAACATCCAGACGCGCTCTCTCGCTACCGCA 822
Db -----
```

Db 232 -----ArgLeuGluArgAsnLeuGlu-----ProAla-----Gln 242
QY 823 AACGTCGAACAGTCGCCCTCTACCAAAATAACAGACGGACGCATCGACTGCAGAG 882
Db 243 GluArgArgLeuProAlaValAlaAlaHisProArgGly-----AlaGln 258
QY 883 CGTCCAAACCGCGCTGAT-----CAGCGACACCCCTGCG 915
Db 259 ArgProGlyProAlaAspProLeuArgArgGlyValLeuArgProGluGlnHisGlnAla 278
QY 916 AAAAGGACTCGA-----CCGCGACCGCGGCAAAACCGCGATTCGCGGAGGCTGCA 966
Db 279 LeuGlyGluArgAlaGlyLeuProArgProArgLeuAlaAsp---ArgProAlaGln 297
QY 967 AGACGCGTCAACAGCCGCAAAAGCGTCTATCGGTTTACCCCTATTCGTCCTCC 1026
Db 298 ProArgAlaAlaAlaArgAlaHisGargAlaGlyAlaGly----- 310
QY 1027 AAAATCCGCGACAGACGACCTGGCAAAACTGGTGCGAGCGCATAGACGTTACCGTCCT 1086
Db 311 -----GluArgGlnGlyProAspGlyGlyArgGlyProAlaAlaAspArgPro 326
QY 1087 GACCACTCGCTACAGCGACCGAGTTCGCGCGCTCAATTCGCGCTACGTCAATACCG 1146
Db 327 ArgProLeuGlnAlaHisGlnArgGln---ProArgProTyrAsnTyrArgHisAlaPro 345
QY 1147 AAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCACCATCGCTCCC 1206
Db 346 GlnGlyGlnGln-----AlaProAlaThrProAlaArg--- 357
QY 1207 GCCCAAAAGACAAAGGCTGACCGGAGCTCGTAACCAAGCTGCATGCCAAACCTT 1266
Db 358 -----ArgAlaLeuProAla-----Val 363
QY 1267 CATGTGACGGCAACCATCTTCATCGCTCAATCAACCTCGACCCCGCTCCGCAAG 1326
Db 364 AlaThrGlyArgArg----- 368
QY 1327 GCTCAATACCAATGGCGTGCATCGAAGCCCAAAATCGCAACAGAT---GGA 1383
Db 369 ---ArgIleArgHisProArgArgGluArgArgProGlySerArgArgAlaProLeuAla 387
QY 1384 CGCAGCCTCGCGATACACACCGAATAGCGCTACCGCTTACCGCTTCGCAAAACACAA 1443
Db 388 AlaHisProArgArgPheGlnArgAlaValArgHisProLeu----- 401
QY 1444 CCGCTGCAATGGCAGCATCCGCGACCCGCAAAACCTTACC-----GAACGAACC 1494
Db 402 ---ProAlaAspLeuTyrGlnArgLysProArgCysGlnProLeuProGlyArgGln 420
QY 1495 CGAAGCCAACTTTGGAACCGCATCGCGCAAAATCCATCCCTGCTGCCCATCGAAGG 1554
Db 421 ArgArgArgProProAspAlaThrArgArgArgAlaValProGlyGlnGlyArg 440

RESULT 34
US-09-252-991A-17231
; Sequence 17231, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17231
; LENGTH: 2294
; TYPE: PR

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17231
Alignment Scores: 1.1e-12 Length: 2294
Pred. No.: 266.50 Matches: 185
Score: 30.6% Conservative: 43
Best Local Similarity: 24.9% Mismatches: 223
Query Match: 9.3% Indels: 293
DB: 2 Gaps: 39
US-10-665-990A-13 (1-1561) x US-09-252-991A-17231 (1-2294)
QY 121 GGAAGCCGTCATTTCAATATCTTCCAAACCTGCTCTCGACAAACATCTCGCAATCCG 180
Db 86 GlyHisProAlaArgTyrHisArgLeuProArgProAlaAlaGluAlaGlyArgHisPro 105
QY 181 GCACACCCCTCATAAACACGGCTATCCGACATCTACCTGCTCTGCACACCC----- 231
Db 106 ArgHisGlyArgGlyArgGlyProArgArgProLeuArgGlnProGlyAlaGly 125
QY 232 -----CCAGAACCCCTTGCCTGCGCCG----- 252
Db 126 SerAlaArgGlnCysArgLeuArgGluProArgProGlyGlyProGlyGlnProGly 145
QY 253 -----CGCGCCCTTATCGAATCTCCGA----- 276
Db 146 ValProAlaAspArgAlaAlaArgProArgArgIleArgArgGlyGlyArgAla 165
QY 277 -----ACACAGCTCGATTTGCAATCTACATTTGCGC 309
Db 166 GlyProArgGluAlaArgGlyProGlyProGlyArgGlyLeuProAlaAlaGly 185
QY 310 CAACGACAT-----TTCCGCGAGGCTGCTGTTCACCTCAT 345
Db 186 AlaArgHisCysArgArgAlaAlaProGlyValLeuArgGlnGlyAlaAlaGlyProAla 205
QY 346 GTACCTGCGCGAGAACG---CGCGCTGCG----- 372
Db 206 ProProAlaArgArgSerProArgArgAlaGlyGlyGlyGlnArgArgGlyArgLeu 225
QY 373 -----CGTACGCTGCTGTGGACGACAAACACACGCGG----- 408
Db 226 GlyAsnArgArgLeuProAlaAlaAlaThrGlyGlnProArgArgGlyProHisTyrSer 245
QY 409 ---GTTGGACATCTCTGCTCGCCT----- 432
Db 246 ProValProArgAlaProAlaAlaProAlaGlyAlaValProArgThrGlyGlyProGly 265
QY 433 -----CGACAGCCATCCCAATATCGAAGTGCCT---GTTCAA 468
Db 266 ArgArgValLeuTyrLeuHisArgGlnProAlaAlaSerArgGlyValProGlyValGln 285
QY 469 CCCTTCTGCTACGCAAAATGCGCGC----- 495
Db 286 ProArgArgProGlyArgGlnLeuArgProGlyAspAlaArgProAlaProGlyAlaGly 305
QY 496 -----ACTCGCTACTGACCGA----- 513
Db 306 AspGlyAspArgArgGlnGlyAspArgHisProAspArgArgGluGlyLeuArgAla 325
QY 514 -----CTTCCCGCTCAACCGCCCGCATGTCACAAATCCTT-----TAC 555
Db 326 ArgArgAspAlaGlnArgProGlnProAspArgGlnArgGlnArgProGlyGlyAla 345
QY 556 CGCGCAACACCGCGCCACCATCTACGCGCGACCAATATCGCGACCAATCTTCAAGT 615
Db 346 ArgArgArgProArgArgGluProAlaAlaArgAlaLeuProAlaAlaGly-----GlnGly 363
QY 616 CGGTGAGGACACCGTTTTCGCCGACCTGGACATCCT----- 651
Db 364 ArgLeuGlyGlyAlaHisProArgProGlyGlnProAlaThrAlaGlyTyrArgProAla 383

```
QY 651 ----- 651
Db 384 AlaAspAlaGlyGlyAlaAlaAspProLeuProArgArgLeuHisLeuPheArgThr 403
QY 652 -----CGCCACGGCAGCGTCT 669
Db 404 GlyProArgGlnArgAlaLeuGluAlaAlaAspAlaLeuArgArgLeuProHis 423
QY 670 CGCGGAAGTATCGCAGCTTCGACCG-----CTACTGGG----- 705
Db 424 ArgArgProValProAlaGluProGlyLeuLeuGlyHisSerArgIleSerAsnAsp 443
QY 706 -----AAGCATTCCGC 717
Db 444 AlaSerHisArgArgSerLeuArgHisProGlyArgArgGlnArgGlyProAlaArg 463
QY 718 CCACAACGCCACGCCATCATCGCAGCGGCAACATCGCAAGGCTCTCAAGCACTCGG 777
Db 464 ArgProHisHisAspHisAlaAlaSerArgArgProArgProGlyThrGlyArgArgPro 483
QY 778 ATACAACGACGAAACATCCAGACACGCGCTCTCGCTACCG-----CGAAAC 825
Db 484 ProArgArgArgSerLeuHisAlaAlaAlaAspGlyGlnProGlyAlaAlaGluArgPro 503
QY 826 CUTCAGAACAGTCCGCCCTCTCAAAAATAATACAGACGGGACGCATCGACTGGCAGAGCGT 885
Db 504 ArgArg-----ArgProGlnProAlaGlyAla--GlyArgArgProThrAlaGlyAsp 520
QY 886 CCNAACCCGCTGATCAGCGACAC-----CCCTGCAAAAGGACT 924
Db 521 AlaHisProProAlaGlnHisHisArgProSerGlyAlaGlyGlnProAlaAlaGluThr 540
QY 925 CGA-----CCCGCAGCCGCGCAA-----ACCGCGATTGCGGAGGCTCGAAGA 969
Db 541 AlaArgLeuProAlaProValArgGlyGluGlyThrProGlyArgArgGlyAlaGlyArg 560
QY 970 CCGGCTCAACAGCCGGAAGAGCGTCTATCTGTTTCACCTATTGCTCCCTACAAA 1029
Db 561 SerAspAlaGlyAlaLeuArg-----AlaLeuHisArgPro----- 572
QY 1030 ATCCGCGCACAGCAGCTCGCAAAATCGTGCA---GGACGCGCATAGAGTTTACCGTCT 1086
Db 573 -----ArgArgSerGlyAlaGlnHisAlaValAlaGlyGlnHisGlnArgLeu----- 587
QY 1087 GACCAACTCGTACAGGCGACGAGCTTGCCGCCCTCCATTCCGCTACGTCAAATACCG 1146
Db 588 GlyGlnAlaGluProAlaAspHisProAlaGlnArgGlyLeuGlyArgArgGluGlyLeu 607
QY 1147 AAAACCGGTGTCAAAGCCGCGCATCAAACTCTACGAGCTGCAACCCACCACTGC----- 1200
Db 608 ProValAla-----ArgThrLeuProAlaGluProAlaProAlaProAla 622
QY 1201 -----CGTCCCGCCACAAAGACAAAGGCTGACCGG 1233
Db 623 ProAlaGlyAlaAlaValProValHisGlnProArgLeuArgArgProLeuProGlyHis 642
QY 1234 CAGCTCCGTAACAGCCT-----GCATGCCAAAACCTTCATTGTGGACGGCAAGC 1284
Db 643 AlaArgArgProGlnProThrGlyGlyAlaAlaArgAlaHisArgGlyGlyAspProGln 662
QY 1285 CATCTTCATCGG-----CTCATTCACCT----- 1308
Db 663 HisProArgArgValArgAlaArgThrValAlaAlaLeuAlaArgProGlnArgAspAla 682
QY 1309 CGACCCCGCTTC-----CGCAGCGGTCAATACCGAAAT----- 1341
Db 683 ArgProProPheAlaIleProAlaAlaValAlaGlyArgTyArgHisArgProGlyLeuAla 702
QY 1342 -----GGCGTCGTCATCGA----- 1356
Db 703 AlaGlyLeuAlaValArgProAlaAspGluAlaGlyArgArgArgAlaGlyValGln 722
QY 1357 ---AAGCCCCAAATCGCAGAACAGTGGAGCGCACCTCGC----- 1395
```

```
Db 723 AlaHisProArgHisArgArgAspProGlyAlaGlyHisArgProGlyGlyAlaAla 742
QY 1396 -----CGATACCACACCGAATACGCTTACCGCTTACCTCGCAACAACACCG 1446
Db 743 AlaGlyAspArgAlaAlaThrProGlyAlaLeuProArg-----ArgArgHisGlnGly 760
QY 1447 CCTGCA-----ATGGCAGATCCCGC-----CACCCGAAAAACCTA 1482
Db 761 ProAlaGlyGlyGlyGlyArgArgArgProLeuGlyGlyAspHisProArgArg 780
QY 1483 CCGNACGAAACCCGAGCCAAACTTTGAAACGATCGCCGCAAAAATCTTATCCCTGCT 1542
Db 781 AlaValArgGlnArgGlnArg-----GlnArgAlaArgValProAlaAlaAla 798
QY 1543 GCCCATCGAAGG 1554
Db 799 AlaHisArgArg 802

RESULT 35
US-09-252-991A-27864
; Sequence 27864, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27864
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27864

Alignment Scores:
Pred. No.: 8,686-13 Length: 777
Score: 265.50 Matches: 158
Percent Similarity: 33.9% Conservative: 49
Best Local Similarity: 25.9% Mismatches: 202
Query Match: 9.3% Indels: 202
DB: 2 Gaps: 31

US-10-665-990A-13 (1-1561) x US-09-252-991A-27864 (1-777)
QY 120 CGGAAGCGCTCATTTCAATA-----CTTCCAACCTGTCC 155
Db 197 ArgLysSerValSerAlaIleHisSerArgGluLeuAlaAspProGluProArgAlaPro 216
QY 156 TCCTGGACAAACATCTCGAAATCCGGACACCCCTCATATA---ACAACGGGCTAT---CCG 209
Db 217 ThrLeuAlaSerSerValArgGlyGlyThrAlaLeuIleArgLeuSerGlyTrpArgPro 236
QY 210 ACATCTACTCTGTCGACGACCCCAACGAG----- 239
Db 237 GlnSerSerValAlaAlaValGlyThrArgAlaSerThrArgAlaGlyGlyAlaThrAla 256
QY 240 ---CCCTTGGCGCGCGCGCGCTTATCGAATCTGCGG---AACACAGCCTCGATTGTC 293
Db 257 TrpProProTrpAlaSerLysAlaSerGlnAlaProSerAsnArgGlnAlaSerCys 276
QY 294 AATACATACATTTGGCGCAACGACATTTCCGGCAGCGCTCTGTCTCA----- 338
Db 277 AsnAspArgGlnGly-----AsnAlaGluAlaGlyCysAlaGluAlaIleThrAspPro 294
QY 339 -----ACCTCATGTACCTTGGCGCAGAACCGCGCGTGCCTGCGTAC 377
```


Db 295 LeuileValMetArgAlaProSerProTyrThrLeuProArgProSerAlaProPheVal 314
Qy 378 GCCTGCTGTGGACACAAACACGCGCGGTTCGACGATCTCTGCTCGCCCTCGACA 437
Db 315 GlyLeuAlaGlyLysThrProThrSerSerArgTrp-ArgArgLeuSerArgGluCysAr 334
Qy 438 GCCATCCCAATATCAAGTGGCGCTT-----GTTCAACCCCTT-----CGTCCTAC 482
Db 334 gGluThrGlyLysGluValSerProGlyCysValAsnLeuLeuAspTrpAsnThrProGl 354
Qy 483 GCAATGGCGGCGCACTCGCTACCTGACCGACTTCCCGCGCTCAACCG----- 531
Db 354 uProProThrArgThrHisAlaThrSerLeuProAlaProProArgArgArgLe 374
Qy 532 -CCGCATGCACAAACAACTTTACCGCGCAACCGCGCACCATATCTCGCGCGACGCA 590
Db 374 uProAspAlaArgArgTrpProGlyGlnArgArgProArgHisArgGlnArgThr 394
Qy 591 ATATCGCGCAGCAATA-----CTTCAAAAGTCGGTGAGGACAC---CGTTT 632
Db 394 r---ArgHisArgLeuGlyProGlyProGlyLeuArgSerProProGlyArgProArgTh 413
Qy 633 TCGCCGACCTGGACAT---CCTCGCCACCGGAGCGTCTGCGGGAATATC-----GC 683
Db 413 rLeuAlaAlaGlyHisArgProProTrpArgGlnArgLeuArgThrGlyThrHisProGl 433
Qy 684 ACGACTTCGACCG-----CTACTGGGCAAGCCATTCCG 716
Db 433 yArgLeuArgProGlyGlyAspArgGlyArgLeuArgArgAlaGlyProGlyAspAs 453
Qy 717 CCCACAGCCGCGCATCATCTCCGACGCGCAACATCGGCAAGGCTTCAAGCACTCG 776
Db 453 pProArgArgGlnAlaGlyArgProProArg----- 463
Qy 777 GATACAACGCAAAACATCCAGACACGCGCTCTCGCTACCGCGAAACCGTCGAACAGT 836
Db 464 ---GlnArgAlaGlyProAspHisArgCysArgProAlaProGlyLleArgArgProGl 482
Qy 837 CGCCCTCTACCAAAATACAGACGGGACGCATCGCTGCGAGAGCGTCCAAACCGCC 896
Db 482 nAlaHisAlaGluSerArgArgSerGlyThrHisArgLeuValGlnArg----- 498
Qy 897 TGATCAGCGACACCCCTGCAAAAGGACTTCGACCGCGACCGCGCAACCGCGATTGCGG 956
Db 499 -----ArgLeuHisProArgArgThrGluAsp----- 507
Qy 957 GGAGCTCAGACGCGCTCAACAGCCCGAAAAAAGCGTCTATCTGTTTCACCCATT 1016
Db 508 -ProAlaArgHisArgAlaHisPro----- 515
Qy 1017 TCGTCCCTACAAATCCGGCACAGACGCACTGGCAAACTGGTCAGGACGGCATAGACG 1076
Db 516 -----HisAspProProGlyGlnCysAlaProGlyArgHisLeuAr 529
Qy 1077 T-----TACCGTCTCT-----GACCAACTCGGTAC 1100
Db 529 gAspProHisProAlaGlyAspHisArgProGlyGluLysProAlaAspGlnProAlaTh 549
Qy 1101 AGCGGACCGA-----CGTTGGCGCGTCCATTCCGGCTA----- 1134
Db 549 rHisHisArgProValProArgAspGlnAlaArgHisProLeuProAlaProArgProGl 569
Qy 1135 -----CGTCAATACGGAACCGTGTCTCA 1160
Db 569 yHisGlyThrProThrGlyGluAspProAlaProGlnArgLeuProArgSerAlaArgAl 589
Qy 1161 AAGCGCGCATCAAACTCTACAGGTGCAACCAACCATGCGCTCCCGCCCAAAAGACA 1220
Db 589 aGlyLleHisProValLeuArgGlyGluInProGlyThrGlnAlaProHisArgHi 609
Qy 1221 AAGGCTGACCGGAGCTCCGTAAACAGCCTGCATGCCAAACCTTATTCGTGACGCGCA 1280
Db 609 sProProGlyProAlaLeuArgGlnArgProAlaLeuArgProAlaGlyCysArgArgGl 629

Qy 1281 AACGCATCTTCATCGG----- 1296
Db 629 nProAspLeuArgArgAspGlyHisArgGlnGlyProAlaProGlyGlyAlaLeuArgLe 649
Qy 1297 -----CTCATTCAACCTCGACC 1313
Db 649 uArgGlyArgSerGlyGlnGluLeuArgAspSerProArgCysGlnArgGlnProArgPr 669
Qy 1314 CCCGTTCCGACGGCTCAATACCGAAATGGCGTCTCATCGAAAGCCCAAAATCGCAG 1373
Db 669 o-----AlaAspProLeuArgArgArgProArg---ArgAr 681
Qy 1374 AACAGATGGAGCGCACCTCGCGATACACACCGGAATACGCTACCGGT----- 1425
Db 681 gAlaGluGlyAla-----SerLeuHisLeuProArgGlyLysGl 694
Qy 1426 -----TACGCTCGACAAACACACCGCCTGCAATGCGACGA-----TC 1463
Db 694 nLeuProSerGlyGlyValProGlnArgArgGlnProAlaValAlaArgArgProGl 714
Qy 1464 CCGCCACCGCAAAACCTACCGGACGAAACCGGAGCCAACTTTGGAAACGCGATCGCG 1523
Db 714 yArgArgAspGlnGlyLeuProArgArg-----ArgHisArgAr 727
Qy 1524 CAAAAATCTATCTCTGCTGCGCCATCGAAGG 1554
Db 727 gProValGlnArgProAlaGlyArgArgArg 737
RESULT 36
US-09-252-991A-27514
; Sequence 27514, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27514
; LENGTH: 627
; TYPE: PR1
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27514
Alignment Scores:
Pred. No.: 9.6e-13 Length: 627
Score: 264.50 Matches: 154
Percent Similarity: 37.0% Conservative: 36
Best Local Similarity: 30.0% Mismatches: 211
Query Match: 9.3% Indels: 113
DB: 2
US-10-665-990A-13 (1-1561) x US-09-252-991A-27514 (1-627)
Qy 112 AGAAGCGGAGGAAAGCGCGTCAATTCATCTTCAA---ACCTGCTCTCTGGACAACAT 168
Db 173 ArgThrAspGlyArgAspArgValArgHisArgLeuArgAlaProProGlyGluGln 192
Qy 169 CCTGCAATCCGCGACACCCCTCATACAACGGGCTATCCGACATCTACCTGCTCGACGA 228
Db 193 Ala-----GluGlyHisProGly-----ProGlyArgGlu 202
Qy 229 CCCCACGAGAGCCCTTCGCGCGCGCGCTTATCGAATCTGCCGAAACACAGCCTCGA 288
Db 203 GlyProGlyHisAlaArgArgSerArgAlaProAlaAlaGlyArgArgArg----- 220

```
QY 289 TTTCGAATACTACTATTGGCGAAGCAGACATTTTCGGCAGGCTGCTTCAACCTCATGTA 348
Db 221 -----HisLeu-----GlnArgArgArgArgAla-----Leu 230
QY 349 CCTTGGCGGAGAACGGCGGTGCG---CGTAGCGCTCTGTTGGACGACACAAACACGCG 405
Db 231 Pro-----ArgIleArgArgAlaAspArgHisPro-----GlyAspProAspProAsp 246
QY 406 CGGTTGGACGATCTCTCTGCTCGCTCGCTCGACGCCATCCCAATAT---CGAAGTGGCGCT 462
Db 247 GlyLeuGlyArgAsnProArgArgProProAlaAspGlyArgHisGlyArgProAlaAsp 266
QY 463 GTTCAACCCCTTCGCTCTACGCAATGGCG-----CGCACTCGGCTACCTACCGACTT 516
Db 267 LeuAlaProLeuArgGlnArgHisProAlaGlyIleArgProGlyAlaArgHisArgGln 286
QY 517 CCCCCGCTCAACCGCGCATGCACAAACATC---CTTTACCGCGCACACCGGCCAC 573
Db 287 ProLeuGlyGlnProProHisArgGlnArgArgGlyLeuTyArgArgSerProLeuHis 306
QY 574 CATACTCGG---CGGACGCAATATCGGCGAAGTACTTTCAAAAGTCGGTAGCACCCT 630
Db 307 ProCysArgHisArgThrAspProAspArgProGlyValHisAlaArgProGlyHisArg 326
QY 631 TTTCGCCGACCTGGACATCTCGCCAC-----CGGCAGCGTCTGTCGG 672
Db 327 LeuArgArgArgArgProArgArgIleProArgSerGlyProArgMetGluGlyArg 346
QY 673 CGAAGTATCGCAGCTTCGACCGCTTACCTGGGCAAGCCATTCGCCCCACCAAGCCAC--- 729
Db 347 ArgGlnAlaGluValAlaArgArgLeuAlaGlyLysLeuProAlaAlaGlnThrHisAla 366
QY 730 ----GGCATCATCGGCGGCGCAACATCGGCAAGGTCTTCAAGCACTCGGTACAACGA 786
Db 367 AlaAlaGlnAspProLeuArgGln---ArgAlaGlyGluThrAlaAlaArgLeuArg 385
QY 787 CGAAATCCAGACAGCGCTCTGCGCTACCGCA-----822
Db 386 AspGluArgValLeuArgGlnGlyCysLeuArgGlnHisArgProValAlaAsp 405
QY 823 AACCGTCGAACAGTGCGCCCT---CTACCACCAAAATATACAGCGGACGCATCGACTGGCA 879
Db 406 ArgArgArgThrValProProCysLeuPro-----ThrAlaProLeuAsp 420
QY 880 GAG---CGTCCAAACCCGCTGATCAGCGACACCCCTGCACAAAGGACT-----924
Db 421 GlnLeuArgProGlyArgProAlaGlyLeuAspHisSerArgArgAlaGlyGlyGln 440
QY 925 ---CGACCGGACCGCGCAACCGCCGATTCGCCGGAG-----960
Db 441 GlyArgProGluProProGlyGlyGlyAlaLeuArgArgLeuArgLeuProValHisAsp 460
QY 961 -----GCTGCAAGACGCGCT---CAACACGCGCGAAAAAGCGTCTATCTGTTTC 1008
Db 461 ArgGlyThrGlyGlyArgArgAlaValGlnAlaAlaLeuTyProArgThrGlyGln 480
QY 1009 ACCCTATTGTCGCC-----TACAAATCCGCGACAGCGACTGGCAAA 1053
Db 481 LeuLeuProGlyProAspProProGlyAlaAlaArgLeuArgHisGlyLeuLeuArgAla 500
QY 1054 ACTGTGACGAGCGGATAGAGTTTACCGTCTCGACCACTCGCTACAGGCGACGAGCT 1113
Db 501 AlaGlyLeuArgGlnHisGlnArgProArgAlaGlyArgLeuArgArgProArgArg 520
QY 1114 TGC CGCGCTCCATTCGCGTACGTCAAATACCGAAACCGCTGCTCAAGCGCGCATCAA 1173
Db 521 GlyGlyArgGlyProGlyLeuGlnGlyAspProGlyVal-----533
QY 1174 ACTTACGAGCTGCAACCAACCATGCGCTCCCGCCACAAAGACAAAGCGCTGACCGG 1233
Db 534 -----ArgProGluArgAspArgArgProGlyGln 544
QY 1234 CAGCTCCGTAACAGCGCTGCATGCCAAACCTTCATTGTCGACGCGCAACGCTTCAT 1293
```

```
Db 545 GlyProArgThrAspAlaThrAlaProGlyAlaGlyGlyArg-----559
QY 1294 CGGCTCATTAACCTCGACCCCGTTCGCGACGGTCAATACCGAAATGGCGTGTGTCAT 1353
Db 560 -----GlyAspProGlyAlaArgHisGlnHisLeuHisGlyTyArgAsp 574
QY 1354 CGAAAGCCCCAAATCGCAGAACAGATGGAGCGGCACCCCTCGCCGA-----TACCAC 1404
Db 575 GlnArgHisGlnArgValArgGlyThrGlyArgGluArgArgArgAlaAspArgHis 594
QY 1405 ACCCGAATACGCCTA-----CGCGTTTACCCTCGACAAACACACCGCT 1449
Db 595 LeuAlaAlaArgLeuIleAlaGlyLeuSerArgArgGlyArgCysValArgHisPro 614
QY 1450 GCAATGGCACGATCCCGCCACCGCAAAACCTTACCCTCGAACGA 1491
Db 615 AlaProGly---GluArgProProGluArgLeuThrGluArg 627

RESULT 37
US-09-252-991A-30396
; Sequence 30396, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30396
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (334)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-30396

Alignment Scores:
Pred. No.: 1,32e-12 Length: 694
Score: 263.00 Matches: 168
Percent Similarity: 34.5% Conservative: 43
Best Local Similarity: 27.5% Mismatches: 216
Query Match: 9.2% Indels: 185
DB: 2 Gaps: 34

US-10-665-990A-13 (1-1561) x US-09-252-991A-30396 (1-694)
QY 111 AAGACGACGACGAAAGCGCTCATTTCAATCTTCAAACTGTCTCTCTGGACACATCC 170
Db 17 ArgArgGlyProAlaAlaArgIleAlaMetValPro-AlaAlaThrProGlyGln-----34
QY 171 TGCAATCCGCGACACACCCCTCATATAACACGGGCT---ATCCGACATCTACCTGCTCGAGC 227
Db 35 -----GlyHisGlyHisArgArgArgGlyGlnLeuArgHis---ArgAlaArgAr 50
QY 228 ACCCGCCAGAGCCCTTGGCCCGCGCGCCCTTATCGAATCTGCCAAACACAGCCTCG 287
Db 50 g-----ProLeuArgPro---ArgProGlyArgArgValArgLeuGlnAr 65
QY 288 ATTTGCATATACTATTGGCGCACACGACATTTCCGCGAGGCTGCTGTTCACCTCATGT 347
Db 65 gValArgLeu-----ArgProAlaAlaArgProAspPr 77
QY 348 ACCTTGGCGGACGCGCGCTGCGGTACGCCT---GCTGTGTCGACGACCAACACGCG 404
```

Db 77 oProGlyValArgHisGlyArgAlaGlnArgProAlaAlaAlaGlyAlaGlnGlyHisG1 97
Qy 405 GCGGTGTGAGATCTCTGCTCGCTCGACAG----- 438
Db 97 yArgLeuProGluSerProArgArgProArgGlnLeuValHisValLeuLeuGlnAr 117
Qy 439 -----CCATCCCAATAT-----CGAAGTCGCGCTGTTCACCC 470
Db 117 gHisLeuArgArgLeuProArgProGlyGlnHisProAlaArgLeuPhePro--ArgPr 136
Qy 471 CTTTCGCTCCTACGAAATGGCGGCGACTCGCTACCTGACCGATTCCTCCCGCGCTCAACC 530
Db 136 oValArgProGluProGlyArgArgLeuLeuHisArgLeuArgLeuArgG1 156
Qy 531 GCCCATGTCACAAATCTTTACCGCGACACCGCGCAC---CATACTCGCGCGAC 587
Db 156 nProAspAlaProAla-----ArgTrpArgProArgSerHisArgArgHisPr 173
Qy 588 GCAATATCGGCGACGAATACTTCAAGTCGGTAGGACACCGT----- 630
Db 173 oGlnProAlaAspGlyValHisProGlyArgHisLeuHisArgArgGlyArgPheProLe 193
Qy 631 -TTTCGCGCGCTCGCATCTCTCCACCGCGACGCTCGTCGGCGAAGTATCGCACGACT 689
Db 193 uValGlnArgAspGlyArgProArg---ProValArgGlyArgHisAlaGlnSerArgCy 212
Qy 690 TCGACCGCTACTGGCGCAAGCCATTC----- 714
Db 212 sArgGlnTrpArgGlyValProValGlyThrAlaAlaLeuProGlnGlyAsnArgArgAs 232
Qy 715 ----CGCCCAACAGCCACGCGCATCATCCGACGCAACATCGGCAAGGGTCTTCAAG 770
Db 232 pAspArgAlaAspArgHisGly-----ArgArgHisArgArgLeuLeuProAs 248
Qy 771 CACTCGGATACAAGA-----CGAATCATCAGACACGCGCTCTCGGCT 815
Db 248 pCysArgProArgArgAspGlnAlaLysHisArgArgLeuProAlaArgProValAlaVa 268
Qy 816 ACCGCGAAACGTCGAACAGTCGCGCCCTCTACCAAAAAATACAGACGGG-----ACGCA 869
Db 268 lArgGlnSerArgArgAlaGlyLeuValArgProProArgGlyGlnAlaAlaLeuAlaHi 288
Qy 870 TCGACTCGCAGAGCGTCAAAACCGCGCTGATCAGCGACACCCCTGCAAAAGAGACTCGACC 929
Db 288 sHisLeuGlyAlaArgProArgSerPro-----ProArgGlySerG1 302
Qy 930 GCGACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGACGCGCTCAAAACAGCCCGAAA 989
Db 302 uArgGlyAlaAlaThrGluHisTrpArgGlyGlnArgLeuArgAlaSerHisGlyLysPr 322
Qy 990 AAAGCGTCTATCGTTTCCACCTATTTCTGCTCCCTACAAATCCGGCACAGCGACTGG 1049
Db 322 oGlyCysAlaAlaGly-----GlyHisProGlyAlaGly***GlyArgGlnGlnG1 339
Qy 1050 CAAACACTGTGTCAGGAGCGCATAGCTTACCGTCT----- 1086
Db 339 y---ThrProAlaGlyHisArgArgArgCysGlnProValArgArgTrpArgProGlyG1 358
Qy 1086 ----- 1086
Db 358 yAlaGlnGlnProAlaSerAlaGlyProGlyLeuLeuCysAspSerGlyAsnLeuGlyHi 378
Qy 1087 -----GACCAACTCGCTACAGCGCGACCGACG 1112
Db 378 sArgAlaValProGlyProProAlaAlaGlyAlaGluProLeuAlaThrGlyGlnArg-- 397
Qy 1113 TTGCGCGCGTCCATTCGCGGTACGTCAAATACCGAAAAACCGCTGCTCAAAAGCGCGATCA 1172
Db 398 ---ArgArgPro-----AlaAlaAlaHisHisProHisAr 408
Qy 1173 AACTCTACGAGTGCACCCCAACCATGCTCCCGGCCACAAAGACAAAGGCGCTGAC-- 1230
Db 408 gPro---GlyAlaAlaArgHisProLeuHisProGlyProArgArgArgLeuProGlyLe 427

Qy 1231 -----CGCGACGTCGTAACACGAGCTGCATGCCAAAACCTTCATTGTGCGCGCAAC 1283
Db 427 uProValAlaArgArgValAlaProAlaAspGlnArgProArgLeuGlyThrAlaAl 447
Qy 1284 GCATCTTCATCGGCTCATTCACACTCGACCC----- 1314
Db 447 aHisAlaThrArg---AlaAspProArgProGlyThrGlyProAlaSerGlyGlyArgLe 466
Qy 1315 -----CCGTCCCGCAGCGCTCAATACCGA----- 1338
Db 466 uSerGlyArgArgSerGlyAlaArgGlnProValAlaThrGlyGlnArgArgLeuG1 486
Qy 1339 -AATGGCGTCGTCGCAAGCCCAAAAT-----CGCAGACACAGATCGAGCGCACCC 1391
Db 486 yAspAlaGlyArgGlnArgHisProGlnHisProAlaGlnArgArgGlyAlaAlaG1 506
Qy 1392 TCCCGCATCACACCGCAATACGCTTACCG---CGTTACCTCCGACAAACACCAACCGCC 1448
Db 506 nArgArgGlyProGlySerArgArgProProGlyArgGlnProGlyGlnArgGlnArgPr 526
Qy 1449 TGCATTTGGCAGCATCC-----CGCCACCGCAAAACCTACCCGAAACGAAACC-- 1494
Db 526 oAlaGlyThrArgGlySerArgAlaArgArgGlnProArgArgArgProArgAlaThrG1 546
Qy 1495 -----CGAAGCGCAAACTTTGGAAACGCAT----- 1518
Db 546 yProLeuAlaThrAlaAlaThrAlaGluAlaArgProAlaLeuArgArgLeuAlaGlyPr 566
Qy 1519 -----CGCCGCAAAAATCTATCCCTGCT 1542
Db 566 oAlaValAlaArgArgIleProProValProAla 577

RESULT 38

US-09-252-991A-30397
; Sequence 30397, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30397
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30397

Alignment Scores:
Pred. No.: 1,59e-12 Length: 894
Score: 262.50 Matches: 173
Percent Similarity: 32.3% Conservative: 47
Best Local Similarity: 25.4% Mismatches: 199
Query Match: 9.2% Indels: 263
DB: 2 Gaps: 40

US-10-665-990A-13 (1-1561) x US-09-252-991A-30397 (1-894)

Qy 160 GGACAAATCTCTCAAAATCCGCGACACCCCTCATACAAACGGGCT----- 204
Db 220 GlyGlnArgProGlnAlaThrArgHisGlyArgAspArgArgAlaProLeuArgLeuAla 239
Qy 205 -----ATCCGACATCTACCTGCTCGAGA-----CCCCCA 234
Db 240 AlaGlyThrProThrGlyProHisArgArgGlnThrProAlaArgArgAlaGlySerPro 259

RESULT 40

US-09-252-991A-25361
; Sequence 25361, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25361
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25361

Alignment Scores:

Pred. No.: 2,02e-12 Length: 1037
Score: 261.50 Matches: 182
Percent Similarity: 33.0% Conservative: 42
Best Local Similarity: 26.8% Mismatches: 191
Query Match: 9.2% Indels: 263
DB: 2 Gaps: 42

US-10-665-990A-13 (1-1561) x US-09-252-991A-25361 (1-1037)

QY	145	CAAACTGTCCTCTGGCAACATCTCGAAATCGCGCACACCCC	-----	189
DB	144	GlnAlaAlaArgProGlyGln---ProAlaGluAspArgProValAlaGlyGluPro	162	
QY	190	---TCATAACAAGGCTATCCGACAT-----CTACCTGCTCGACGACCCCA	234	
DB	163	GlyAspArgGlnArgGlnSerHisGlyAlaAlaGluLeuProTrpArgAspProGlu	182	
QY	235	CGAAGCCCTTGC---CGCCGCGCGCCCTTATCGAATCTGCCGAACA-----	279	
DB	183	GluLeuProPheProArgAlaArgThrGlyArgArgLeuArgArgSerHisProProAla	202	
QY	280	-----CAGCTCGATTGCAATACATATTGGCGCAACGACATTTCCGGCAGGCT	330	
DB	203	GlnProGluArgProGlnSerSerHisGlnArgLeuArgGlnArgHis-----	218	
QY	331	GCTGTTCAACCTCATGCTTCCGCGCAGAACGCGCGTGGCGTACGCTGCTGTGGA	390	
DB	219	-----GlnGlyArgTyrArgArgSerArgLysProAla-----	230	
QY	391	CGACAAACAACGCGCGGTGGACCATCTCTGCTCGCCCTCGA-----	435	
DB	231	-----HisArgLeuAlaGlyArgGlnLeuGlnArgProArgGlnProGlyAsp	247	
QY	436	-----CAGCCATCCCAATATCATGAGTGCGCTGTT	465	
DB	248	AlaLeuGlnArgProGlyGlyTyrLeuGluProAlaLeuProArgProGlnProAla	267	
QY	466	-----CAACCCCTTCTGCTACGAAATGGCGCGCACTCGGTACTCGACCGA	513	
DB	268	ValAlaAlaPheGlnProIleAspProAspAlaLeuAla---AlaLeuLeuGlnAspPro	286	
QY	514	CTTCCCGCCCTCAACCG-----CGCATGCAACAACAAATCCTTTACCGCGCA	561	
DB	287	ArgProGlyProGlnProAlaArgProAspArgGlnArgProLeuGlnArgAlaAsp	306	
QY	562	-----CAACCG-----	567	
DB	307	ProGlyArgArgProValAlaGluProGlnProAlaLysGlyArgTrpArgHisPro	326	
QY	568	-----CGCACCATACT	579	

DB	327	ValArgLeuHisArgAlaGluArgProAspLeuProLeuArgLysGlnArgHisAlaVal	346	
QY	580	CGCGGACGCAATATCGCGCAGCAATA-----	606	
DB	347	ArgArgLeuGlnArgProArgArgAlaLeuValGluAspProValProGlyGlnProGly	366	
QY	607	CTTCAAAGTCGGTAGAGACACCGTTTTCGCGCA-----	639	
DB	367	LeuProArgProGlyLeuAlaValArgArgLeuArgProTrpProGlyArgPro	386	
QY	640	---CCTGGACATCTCCCGCACCGCAGCGTCT-----	669	
DB	387	AspProGlyArgProGlyGlnArgGlyLeuArgLeuProLeuGlnProGlnTrpGlnGlu	406	
QY	670	-----CGCGAAGTATCGCACACTTCCA-----	693	
DB	407	ArgProAlaLeuGlyThrArgProValAlaAlaLeuArgLeuProArgArgProGlyGln	426	
QY	694	-----CCGCTACTGGCAAGCCATTCCGCCCAACAGCCACCGCATCAT	738	
DB	427	GlyProGlnArgAspProAlaLeuGlyHisPro---ProProArgArgArgLeuHisArg	445	
QY	739	CGCAGCGCAACATCGCGCAAGGCTCTCAAGCACTCGGATACCAACAGCAACATCCAG	798	
DB	446	ProArgGlnHisProArgGln-----LeuGlnLeuArg-----	456	
QY	799	ACACGCGTCTCGCTACCGCAACCGTCGAACAGTCGCGCCCTCTACCAAAAATACA	858	
DB	457	---ArgIleProGlyGlyArgLeuProAspGlnValAlaLeuAsnProArgAsn---	474	
QY	859	GACGGAGCATCGACTGCGAGCGCTCAAAACCGCCTGATCAG---CGACACCCCTGC	915	
DB	475	GlnLysLysGluLysLeuProHisAlaCysSerProAlaThrHisLeuArgArgProLeu	494	
QY	916	AAAGGACTCGACCGCGCAACCGCGGATTCGCGGGAGGCTGCAAGACGCGCT	975	
DB	495	ArgArgThrLeuProGlyGlyProArg-----ArgArgArgAlaAlaAla	510	
QY	976	CAACAGCCCGAAAAAGGCTATCTGCTTTCACCCCTATTTTCCTTACAAATCCCG	1035	
DB	511	AlaGlyAlaArgGlnArgGlnHisGly-----ArgArgGlnProSerArg	526	
QY	1036	CACAGA---CGCAGCTGC-----AAACTGTCGAGGACGCATAGAGTTACCGT	1083	
DB	527	HisAlaLeuArgThrAlaProAlaIleArgSerGlyLysGlnArgGlnArgArgArg	546	
QY	1084	CCT---GACCAACTCGCTACAGC-----	1104	
DB	547	AlaPheGlySerArgSerAlaThrGlySerTrpHisAlaValAlaAlaLeuArgLeu	566	
QY	1105	-----GACCGAGTTGCGCGTCCATTCGCGTACGTCAATACCGAAACCGCT	1155	
DB	567	ProGlnProGluHisProAlaArgProLeuArgArgArg-----SerAla	582	
QY	1156	GCTCAAGCCGCACTCAAACTCTAGAGCTGCAACCCACCATCGCTCCCGCCACAAA	1215	
DB	583	AlaAlaAlaGlyThrProGlyProArgTyrArgArgArgProGlyArgProAlaHisPro	602	
QY	1216	AGACAAAGCCCTGACCGCGAG-----	1236	
DB	603	ProGlyArgProAspArgGlnSerAlaAspAlaAlaMetGlyAlaArgTyrAlaValArg	622	
QY	1237	-----CTCGTAACCGCTCGATGCCAAACCTTCATTGTT	1272	
DB	623	ProProLeuAlaAlaProGluLeuArgLeuProAlaHisArgArgLeuArgGln	642	
QY	1273	GGA-----CGCAACCGCATCTTCATCGCTCATCAACCT	1308	
DB	643	GlyArgArgAspGlnHisArgGlnGlnHisLeuAsp-----LeuGlnProAlaLeuArg	660	
QY	1309	-----CGACCCCGCTTCGCGACGCTCAA	1332	

```
Db      661 LeuHisLeuGlyThrGlyArgAlaProGlyThrGlnArgProPro-----Ala 676
QY      1333 TACCGAAATGGCGTCGTCATCGAAAGCCC-----CAAATCGCAGACCA 1377
Db      677 LeuArgLeuAspGlnProGlnArgProLeuAlaProGlyArgLysArgArgThr 696
QY      1378 GATGGA-----GCGCACCCCTCGCCGATAC 1401
Db      697 AlaGlyThrGlyAlaAlaArgGluLeuLeuAspLeuLeuArgAlaGlyProAlaLeuAla 716
QY      1402 CACACCCGAATACGCTACCG-----CGTTACCTCGACAAACACAAACCG 1446
Db      717 AspArgProGlyArgLeuProValGluThrAspGlnArgProHisArg---ArgPro 735
QY      1447 CCTGCA-----ATGGCAGCATCC-----CGCCACCCGAAAAAC-----CTACCC 1485
Db      736 ProAlaGlyArgLeuAlaArgAlaGlyLeuArgHisArgProGlyGlyAspValGlnPro 755
QY      1486 GAACGAACCCGAAGC-----CAAACCTTTGAAACGCGATCGCCGCCAAAAATCCT 1533
Db      756 GlyGlnAlaAsnAlaValArgGlnProLeuArgLysArgArgAlaGlnPro 773
```

Search completed: May 2, 2006, 05:28:41
Job time : 146.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:01:11 ; Search time 51.6 Seconds
(without alignments)
2658.412 Million cell updates/sec

Title: US-10-665-990A-13

Perfect score: 2852

Sequence: 1 caaatacaggaatgcgcgt.....tgcccatcgagggtttatta 1561

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB spool/US10665990/runat_01052006_111938_9478/app_query.fasta_1
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOCPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10665990 @CGN 1 1 348 @runat_01052006_111938_9478 -NCPU=6 -ICPU=3
-NO_MMAP -NRG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq 21:.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2681	94.0	525	6	ABP79466 N. gonorr
2	2630	92.2	525	5	ABB78070 Amino aci
3	2606.5	91.4	507	3	AAY75751 Neisseria
4	2586	90.7	508	3	AAY75753 Neisseria
5	2573	90.2	508	3	AAY75752 Neisseria
6	831.5	29.2	543	8	ADL04719 M. catarr
7	790	27.7	564	7	ABO75163 Pseudomon
8	632	22.2	502	2	AAW55452 H. pylori
9	632	22.2	502	2	AAW55550 H. pylori

10	590	20.7	428	2	AAW98378 H. pylori
11	493.5	17.3	271	6	ADA34272 Acinetoba
12	478.5	16.8	275	6	ADA34304 Acinetoba
13	421	14.8	2519	7	ABM88218 Rice abio
14	409	14.3	321	6	ADA36398 Acinetoba
15	336.5	11.8	5178	9	AEA04533 Human pro
16	336.5	11.8	5179	6	AAW24516 C899p pre
17	336.5	11.8	5179	6	ABP55365 Human col
18	336.5	11.8	5179	6	ABO07258 Human p53
19	336.5	11.8	5179	7	ADD48091 Human pro
20	336.5	11.8	5179	7	ADD44998 Human pro
21	336.5	11.8	5179	8	ADQ29695 Human col
22	336.5	11.8	5179	8	ADQ80379 Intestina
23	336.5	11.8	5179	9	AEA27639 Human int
24	328	11.5	395	4	ABU53160 Human tes
25	326.5	11.4	822	7	ABO73174 Pseudomon
26	325	11.4	386	4	ABU53159 Human tes
27	325	11.4	717	4	ABU53144 Human tes
28	318	11.2	385	4	ABU53156 Human tes
29	314.5	11.0	387	4	ABU53161 Human tes
30	310.5	10.9	368	4	ABU53157 Human tes
31	304	10.7	513	6	ABM71816 Staphyloc
32	303.5	10.6	956	6	ABO53053 Human put
33	303.5	10.6	2296	8	ADN60280 Human spl
34	303.5	10.6	2752	8	ADN60278 Human ser
35	303.5	10.6	2752	8	ADP24102 PRO polyd
36	303.5	10.6	2752	9	ADY91619 Human pro
37	303	10.6	655	7	ABO68366 Pseudomon
38	302	10.6	618	7	ABO74950 Pseudomon
39	301.5	10.6	2263	7	ADJ70425 Human hea
40	301	10.6	618	7	ABO79612 Pseudomon
41	300.5	10.5	695	4	ABU53152 Human tes
42	300.5	10.5	745	4	ABU53154 Human tes
43	296	10.4	486	7	ADG30702 Xanthomon
44	295	10.3	507	7	ABO74852 Pseudomon
45	294.5	10.3	504	5	ABB49813 Listeria

ALIGNMENTS

RESULT 1
ABP79466
ID ABP79466 standard; protein; 525 AA.
XX
AC ABP79466;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 5462.
XX
KW Antibacterial; infection; vaccine; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
FI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
WI; 2003-058415/05.
XX
DR N-PSDB; ABZ40436.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX
PS medicament for treating or preventing N. gonorrhoeae infection.
PS
Disclosure; Page 587; 815pp; English.


```

XX 31-JAN-2001; 2001US-0266070P.
PR 06-AUG-2001; 2001US-0310356P.
PR 23-OCT-2001; 2001US-0344452P.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
PA (APIC ) APICELLA M A.
PA (EDWA ) EDWARDS J L.
PA (GIBS ) GIBSON B W.
PA (SCHE ) SCHEFFLER K.
PA (BROW ) BROWN E.
XX
PI Apicella MA, Edwards JL, Gibson BW, Scheffler K, Brown E;
XX WPI; 2002-619227/66.
DR N-PSDB; ABQ78301.
XX
PT New polypeptide comprising p177, p88, p64, p55 or p46 from Neisseria
PT gonorrhea, useful for preventing, or protecting a female patient
PT against, N. gonorrhea colonization or infection.
XX
PS Claim 7; Page 120-121; 130pp; English.
XX
CC The present sequence represents a p55 polypeptide. The specification
CC describes p177, p88, p64, p55 and p46 polypeptides from Neisseria
CC gonorrhea. The polypeptides are useful as vaccines, for preventing, or
CC protecting a female patient against, N. gonorrhea colonization or
CC infection. Such immunisation can prevent gonorrhea in women. (Updated on
CC 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 525 AA;

```

Alignment Scores:

```

Pred No.: 5,4e-185 Length: 525
Score: 2630.00 Matches: 509
Percent Similarity: 98.7% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 7
Query Match: 92.2% Indels: 0
DB: 5 Gaps: 0

```

US-10-665-990A-13 (1-1561) x ABB78070 (1-525)

```

Qy 2 AAAATACAGCAATGCCGCTCTGAAACTATATCCCGGATGAAAAACACGAGCCTCATTTCC 61
Dy |||||
Dy 6 LysThrGlnAlaMetProSerGluThrIleSerLeuMetLysThrArgSerLeuIleSer 25
Qy 62 CTTTATGCTCTCTCTGTTCAATGTTCTTCATGTTGCCCCCACCCTGGAGAACGAGC 121
Dy |||||
Dy 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProLeuGluGluArgThr 45
Qy 122 GAAAGCGTCAATTTCAATCTTCAAAACCTGCTCTCGGACACATCTCGCAATCGG 181
Dy |||||
Dy 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
Qy 182 CACACCCCTCATAAACAGGCTATCGACATCTACCTGCTCGACGACGCCCCCAGAGCC 241
Dy |||||
Dy 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
Qy 242 CTTGCCGCGCGCGGCTTATCGAATCTGCGGAAACACAGCCTCGATTGCAATATAC 301
Dy |||||
Dy 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyr 105
Qy 302 ATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCACCTCATGTACCTTGGCGAGAA 361
Dy |||||
Dy 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
Qy 362 CGCGCGTGGCGGTACGCTCTTGGACGACAAACACACGCGCGGTTGGACGATCTC 421
Dy |||||
Dy 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
Qy 422 CTGCTCGCCTCGACAGCATTCCTCAATATCGAAGTGGCCTGTTCACCCCTTCGTCCTA 481
Dy |||||

```

```

Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
Qy 482 CGCAAAATGGCGCGCACTCGGCTACTGACCGACTTCCCCCGCTCAACCGCGCATGCAC 541
Dy |||||
Dy 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMethHis 185
Qy 542 AACAAATCCTTTACCGCGCAACACCGCGCACTACTCGGCGGACGCAATATCGCGCAC 601
Dy |||||
Dy 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyArgAsnIleGlyAsp 205
Qy 602 GAATACTTCAAAGTCGTGAGGACACCCGTTTCCCGCACTCGGACATCTCGCCACCGCGC 661
Dy |||||
Dy 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
Qy 662 AGCGTCGTCGCGGAAGTATCGCAGCACTTCGACCGCTACTGGGCAAGCCATTCGCCCCAC 721
Dy |||||
Dy 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
Qy 722 AACGCCACGCGCATCATCCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATAC 781
Dy |||||
Dy 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
Qy 782 AACGACGAAACATCCAGACGCGCTCTCTGCGCTACCGCGAAACCGTCGAAACAGTCGCC 841
Dy |||||
Dy 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
Qy 842 CTCTACCAAAAATACAGAGCGGACGATCGACTGCGGACGCGTCCAAAACCGCGCTGATC 901
Dy |||||
Dy 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
Qy 902 AGCGACACCCCTTGCAAAAGGACTCGACCGGACCGCGCAACCGCGATTCGCGGGAGG 961
Dy |||||
Dy 306 SerAspAspProAlaLysGlyLeuAspArgAspArgLysProProlIleAlaGlyArg 325
Qy 962 CTGCAAGACGCGCTCAAAACGCGCAAAAAGCGTCTATCTGTTTCACTTATTCGTC 1021
Dy |||||
Dy 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
Qy 1022 CCTACAAAATCCGCGACAGACGCTGGCAAACTGCTGAGGACGCGATAGAGTTACC 1081
Dy |||||
Dy 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
Qy 1082 GTCTGACCAACTCGCTTACAGGCGACCGCGTTCGCGCGCTCCATTCGCGTACGTCAAA 1141
Dy |||||
Dy 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
Qy 1142 TACGAAAACCGCTGCTCAAAAGCGCGCATCAAACTCTACAGCTGCAACCCCAACATGCC 1201
Dy |||||
Dy 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
Qy 1202 GTCCCGCCCAAAAGACAAAGCGCTGACCGGAGCTCCGTAACCGCTGCGATGCCAAA 1261
Dy |||||
Dy 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
Qy 1262 ACCTTCATTGTGACGCGCAAAACGATCTTCATCGGCTCATTCACCTCGACCCCGCTTCC 1321
Dy |||||
Dy 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
Qy 1322 GCAAGGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAATTCGCAGAACAGATG 1381
Dy |||||
Dy 446 AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
Qy 1382 GAGCGACCCCTCCCGATACCCACCGATACCGCTACCGCTTACCTCGCAACACAC 1441
Dy |||||
Dy 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
Qy 1442 AACCGCTGCAATGGCAGCATCCCGCACCCGCAAAAACCTTACCGAACGCAACCCGAGGCC 1501
Dy |||||
Dy 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
Qy 1502 AAACTTTGGAAACGCAATCGCCCGCAAAATCTATCTGCTGCTGCGCAATCGAAGGTTTATTA 1561
Dy |||||
Dy 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525

```



```
Db      201 AspLeuAspIleLeuAlaThrGlySerValValGlyGluValSerHisAspPheAspArg 220
      698 TACTGGCAAGCATTCCGCCACACACGCGCATCATCCGACGGGCAACATCGGC 757
      221 TyrTrpAlaSerHisSerAlaHisAsnAlaThrArgIleIleArgSerGlyAsnIleGly 240
      758 AAGGTCCTTCAAGCACTCGATACAAACGACGAAATCATCCAGACGCGCTCTCGCTAC 817
      241 LysGlyLeuGlnAlaLeuGlyTyrAsnAspGluThrSerArgHisAlaLeuLeuArgTyr 260
      818 CCGCAAAACCGTCGACAGTCGCCCTCTACCAAAAATACACAGCGGACGATCGACTGG 877
      261 ArgGluThrValGlnGlnSerProLeuTyrGlnLysIleGlnThrGlyArgIleAspIrp 280
      878 CAGAGCGTCAAAACCGCTGTATGACGACACCCCTGCAGAAAGACTCGACGCGACGCG 937
      281 GlnSerValGlnThrArgLeuIleSerAspAspProAlaLysGlyLeuAspArgAspArg 300
      938 CCAAAACCGCGATTCCCGGAGGTGCAAGCGGCTCAACACAGCCGCAAAAAGCGTC 997
      301 ArgLysProProIleAlaGlyArgLeuGlnAspAlaLeuLysGlnProGluLysSerVal 320
      998 TATCTGGTTTCACCTATTTCCTCACAATCCGGCACAGCGCACTGCGCAAACTG 1057
      321 TyrLeuValSerProTyrPheValProThrLysSerGlyThrAspAlaLeuAlaLysLeu 340
      1058 GTGACGAGCGCATAGACGTTTACCGTCTGACCAACTCGCTACAGCGCACCGACGTTGCC 1117
      341 ValGlnAspGlyIleAspValThrValLeuThrAsnSerLeuGlnAlaThrAspValAla 360
      1118 GCGGTCCATTCCGGCTACGTCAAATACGAAACCGCTGCTCAAGCGGGCATCAAACTC 1177
      361 AlaValHisSerGlyTyrValLysTyrArgLysProLeuLeuLysAlaGlyIleLysLeu 380
      1178 TAGGAGTCCAACCAACCATCCGTCGCCGCCACAAAAGCAAGGCTGACCGGCAGC 1237
      381 TyrGluLeuGlnProAsnHisAlaValProAlaThrLysAspLysGlyLeuThrGlySer 400
      1238 TCCGTAACACGAGCTGCATGCCAAAACCTTCATTGTGGACGGCAAAACGCATCTTCATCGC 1297
      401 SerValThrSerLeuHisAlaLysThrPheIleValAspGlyLysArgIlePheIleGly 420
      1298 TCATTCAACTGACCCCGTTCGACGCGCTCAATACCGAAATGGCGTCTCATCGAA 1357
      421 SerPheAsnLeuAspProArgSerAlaArgLeuAsnThrGluMetGlyValValIleGlu 440
      1358 AGCCCCAAAATCGCAGAACAGATGGAGCGCACCCCTCGCGGATACACACCCGCAATACGCC 1417
      441 SerProLysIleAlaGluGlnMetGluArgThrLeuAlaAspThrSerProGluTyrAla 460
      1418 TACCGGTTTACCTCGACAAACACACCCCGTGCATGTCAGATCCCGCCACCCGAAAA 1477
      461 TyrArgValThrLeuAspArgHisAsnArgLeuGlnTrpHisAspProAlaThrArgLys 480
      1478 ACTACCCCAAGCAAGCCCAAGCCAACTTCGGAACGCATCGCGCAAAATCTCATCC 1537
      481 ThrTyrProAsnGluProGluAlaLysLeuTrpLysArgIleAlaAlaLysIleLeuSer 500
      1538 CTGTCGCCCATCGAAGGTTTATTA 1561
      501 LeuLeuProIleGluSerLeuLeu 508

RESULT 5
ID AAY75752
XX AAY75752 standard; protein; 508 AA.
AC AAY75752;
XX
XX 21-MAR-2000 (first entry)
XX
XX Neisseria meningitidis ORF 987 protein sequence SEQ ID NO:2976.
XX
```

```
KW      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW      antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW      antibacterial; gene therapy.
XX
OS      Neisseria meningitidis.
PN      WO9957280-A2.
XX
PD      11-NOV-1999.
XX
PF      30-APR-1999; 99WO-US009346.
XX
PR      01-MAY-1998; 98US-0083758P.
PR      31-JUL-1998; 98US-0094869P.
PR      02-SEP-1998; 98US-0098994P.
PR      02-SEP-1998; 98US-0099062P.
PR      09-OCT-1998; 98US-0103749P.
PR      09-OCT-1998; 98US-0103794P.
PR      09-OCT-1998; 98US-0103796P.
PR      25-FEB-1999; 99US-0121528P.
XX
PA      (CHIR ) CHIRON CORP.
PA      (GENO-) INST GENOMIC RES.
XX
XX      Frazer C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI      Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI      Tettelin H, Venter JC;
XX
DR      WPI; 2000-062150/05.
DR      N-PSDB; AAZ54514.
XX
XX      Novel Neisserial polypeptides predicted to be useful antigens for
PT      vaccines and diagnostics.
PT
PS      Claim 2; Page 1395; 1453pp; English.
XX
XX      AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAY75941
CC      represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC      and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC      PCR primers used in the exemplification of the present invention. The
CC      polypeptides, the polynucleotides, antibodies and compositions of the
CC      invention can be used as vaccines, as diagnostic reagents, and as
CC      immunogenic compositions. The polypeptides can be used in the manufacture
CC      of medicaments for treating or preventing infection due to Neisserial
CC      bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC      Neisseria bacteria, or to raise antibodies. They may also be used to
CC      screen for agonists or antagonists, which may themselves have use as
CC      antibacterial agents. The polynucleotides of the invention may also be
CC      used in gene therapy protocols
XX
SQ      Sequence 508 AA;

Alignment Scores:
Pred. No.:      8,46e-181      Length:      508
Score:          2573.00      Matches:      497
Percent Similarity: 98.8%      Conservative: 5
Best Local Similarity: 97.8%      Mismatches: 6
Query Match:    90.2%      Indels:      0
DB:            3      Gaps:      0

US-10-665-990A-13 (1-1561) x AAY75752 (1-508)
QY      38 ATGAAAAACGACGAGCCTCATTTCCCTTTTATGCTCTCTCTGTCATGTTCTTCATGG 97
Db      1 MetLysThrArgSerLeuIleSerLeuLeuCysLeuLeuLeuCysSerCysSerIrp 20
QY      98 TTGCCCCCAGCTGGAAGAACGACGAGAACGCGCTCATTTCAATACATTCCAAACCTGTCTC 157
Db      21 LeuProProLeuGluGluArgThrGluSerArgHisPheAsnThrSerLysProValArg 40
QY      158 CTGACAAATCTCTGCAAAATCCGGCACACCCCTCATATAACAACGGGCTATCCGACATCTAC 217
Db      41 LeuAspAsnIleLeuGlnIleArgHisThrProHisThrAsnGlyLeuSerAspIleTyr 60
```

```
QY 218 CTGCTCGAGACCCCAAGCCCTTGCGCGCCCGCGCCCTTATCGAATCTCCCA 277
DB : : : : :
DB 61 LeuLeuAsnAspProHisGluAlaPheAlaAlaArgAlaAlaLeuLeuGluSerAlaGlu 80
QY 278 CACAGCTCGATTTCGAATACATATTGGCGCAACGACATTTCCGGCAGGCTGTGTC 337
DB : : : : :
DB 81 HisSerLeuAspLeuGlnTyrTyrIleTrpArgAsnAspIleSerGlyArgLeuLeuPhe 100
QY 338 AACCTCATGTACTTGGCGCAGAACGCGCGTGGCGGTACGCTCTGTCGAGACACAC 397
DB : : : : :
DB 101 AsnLeuValTyrLeuAlaAlaGluArgGlyValArgValArgLeuLeuAspAsn 120
QY 398 AACCGCGCGGTTCGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCAAGTG 457
DB : : : : :
DB 121 AsnThrArgGlyLeuAspAspLeuLeuAlaLeuAspSerHisProAsnIleGluVal 140
QY 458 CGCTGTTCACCCCTTGCTCTAGCAAAATGGCGCGCACTCGGCTACTGACCGACTTC 517
DB : : : : :
DB 141 ArgLeuPheAsnProPheValLeuArgLysTrpArgAlaLeuGlyTyrLeuThrAspPhe 160
QY 518 CCCGCTCAACCGCGCATGCACAAACAAATCTTTACCGCGCACACCGCGCCACCAT 577
DB : : : : :
DB 161 ProArgLeuAsnArgGlyMetHisAsnLysSerPheThrAlaAspAsnArgAlaThrIle 180
QY 578 CTCGCGGACGCAATATCGCGACGAATACCTTCAAAGTCGCTGAGCACACCGCTTTTCGCC 637
DB : : : : :
DB 181 LeuGlyGlyArgAsnIleGlyAspGluTyrPheLysValGlyGluAspThrValPheAla 200
QY 638 GACCTGGACATCTCGCCACCGGAGCGTCTCGCGCAAGTATCGCAGACTTCGACCGC 697
DB : : : : :
DB 201 AspLeuAspIleLeuAlaThrGlySerValValGlyGluValSerHisAspPheAspArg 220
QY 698 TACTGGCAAGCATTCCGCGCCACAAACCGCGCATCATCCGCGCGCAACATCGGC 757
DB : : : : :
DB 221 TyrTrpAlaSerHisSerAlaHisAsnAlaThrArgIleIleArgSerGlyAspIleGly 240
QY 758 AAGGCTCTTAAGCACTCGGATACAAACGACGAAACATCCAGACACGCGCTCCTCGCTAC 817
DB : : : : :
DB 241 LysGlyLeuGlnAlaLeuGlyTyrAsnAspGluThrSerArgHisAlaLeuLeuArgTyr 260
QY 818 CGCGAAGCTCGAACAGTGCCTCTACCAAAATAACAGCGGACGATCGACTGG 877
DB : : : : :
DB 261 ArgGluThrValGluGlnSerProLeuTyrGlnLysIleGlnThrGlyCysIleAspTrp 280
QY 878 CAGAGCTCCAAACCGCGCTCATGACGACACCCCTGCAGAAAGACTCGACCGGACCGC 937
DB : : : : :
DB 281 GlnSerValArgThrArgLeuIleSerAspAspProAlaLysGlyLeuAspArgAspArg 300
QY 938 CGCAAAACCGCGATTTCGCGGAGGCTGCAAGACGCGCTCAAAACAGCCGCAAAAGGCTC 997
DB : : : : :
DB 301 ArgLysProProIleAlaGlyArgLeuGlnAspAlaLeuLysGlnProGluLysSerVal 320
QY 998 TATCTGGTTTCACTTTCCTTTCCTTACAAATTCGCGCAACAGCACTGGCAAAACTG 1057
DB : : : : :
DB 321 TyrLeuValSerProTyrPheValProThrLysSerGlyThrAspAlaLeuAlaLysLeu 340
QY 1058 GTGAGGACGCGCATAGAGTTACGCTCTGACCACTCGCTACAGGCGCAGCGTGGCC 1117
DB : : : : :
DB 341 ValGlnAspGlyIleAspValThrValLeuThrAsnSerLeuGlnAlaThrAspValAla 360
QY 1118 GCCGTTCATTCGGCTACGTCAAATACCGAAACCGCTGCTCAAAAGCGCGCATCAAACTC 1177
DB : : : : :
DB 361 AlaValHisSerGlyTyrValLysTyrArgLysProLeuLeuLysAlaGlyIleLysLeu 380
QY 1178 TACAGCTGCAACCAACCATTCGCTCCCGCCCAAAAGACAAAGCGCTGACCGGACG 1237
DB : : : : :
DB 381 TyrGluLeuGlnProAsnHisAlaValProAlaThrLysAspLysGlyLeuThrGlySer 400
QY 1238 TCCGTAAACGCTGCATGCGCAAACTTCATCTGGACGCGCAACCGCATCTTCATCGGC 1297
DB : : : : :
DB 401 SerValThrSerLeuHisAlaLysThrPheIleValAspGlyLysArgIlePheIleGly 420
```

```
QY 1298 TCATTCACTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGCGTCTGTCATCGAA 1357
DB : : : : :
DB 421 SerPheAsnLeuAspProArgSerAlaArgLeuAsnThrGluMetGlyValIleGlu 440
QY 1358 AGCCCCAAATTCGAGAACAGATGGAGCGCACCTTCGCCGATACACACCGGAATACGCC 1417
DB : : : : :
DB 441 SerProLysIleAlaGluGlnMetGluArgThrLeuAlaAspThrThrProAlaTyrAla 460
QY 1418 TACCGCTTACCTCGACAAACACACACCGCTGCAATGGCAGCATCCCGCCACCGCAAAA 1477
DB : : : : :
DB 461 TyrArgValThrLeuAspArgHisAsnArgLeuGlnTrpHisAspProAlaThrArgLys 480
QY 1478 ACTTACCGCAAGAACCGCAAGCAAACTTTGGAAACCATCGCGCAAAATCCTATCC 1537
DB : : : : :
DB 481 ThrTyrProAsnGluProGluAlaLysLeuTrpLysArgIleAlaAlaLysIleLeuSer 500
QY 1538 CTGCTGCCCATCGAAGGTTTATTA 1561
DB : : : : :
DB 501 LeuLeuProIleGluGlyLeuLeu 508
RESULT 6
ADL04719
ID ADL04719 standard; protein; 543 AA.
XX
AC ADL04719;
XX
DT 06-MAY-2004 (first entry)
XX
DE M. catarrhalis protein #485.
XX
KW Moraxella catarrhalis; infection.
XX
OS Moraxella catarrhalis.
XX
FN US6673910-B1.
XX
PD 06-JAN-2004.
XX
PF 04-APR-2000; 2000US-00540236.
XX
PR 08-APR-1999; 99US-0128416P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2004-178127/17.
DR N-PSDB; ADL02799.
XX
PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.
XX
PS Disclosure; SEQ ID NO 2405; 429pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents the amino acid
CC sequence of a M. catarrhalis protein.
XX
SQ Sequence 543 AA;
Alignment Scores:
Pred. No.: 1.62e-52 Length: 543
Score: 831.50 Matches: 194
Percent Similarity: 54.7% Conservative: 102
Best Local Similarity: 35.9% Mismatches: 198
Query Match: 29.2% Indels: 47
DB: 8 Gaps: 12
US-10-665-990A-13 (1-1561) x ADL04719 (1-543)
```

```
QY 50 AGCCTCATTTCCCTTTTATGCTCCTCTCTCTGTTTCATGTTCTTCTATGTTGTCGCC----- 103
Db 13 SerIleHisThrValMetLeuValLeuValGlyCysLysAsn---LeuProAspThr 31
QY 104 -----CCATCTGGAAGAACGGAGGAAAGCGCTCATTTCAATACT 142
Db 32 ProHisLeuProLysSerLeuAlaLeuThrLysGlnValGlnGluArgHisGlnMetVal 51
QY 143 TCCAAACCTGCTCCTCTGGACAAACATCTCTG-----CAAATCCGGCACACC 187
Db 52 LysHisGlyGlySerSerAspGlyLeuIleSerAlaIleAsnAlaGlnSerHisHis 71
QY 188 CTTCTATAACAACGGGTATCCGACATCTACCTGCTCGAGGACCCCAAGACGCTTGGC 247
Db 72 ProAsnGlnSerGly-----TyrTyrProIleThrThrGlyAlaAsnAlaPheAla 88
QY 248 GCCCGCGCGCTTATCGAATCTCCGAACACAGCTCGATTTCGAATACATACATTTGG 307
Db 89 AlaArgSerThrLeuThrAspValAlaAsnGlnSerIleAspIleGlnTyrTyrIleTrp 108
QY 308 CGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGTACTCTTCCGCGAGAAGCGGC 367
Db 109 HisAsnAspGluAlaGlyGlnLeuMetLeuLysAspLeuTrpGluAlaAlaAspArgGly 128
QY 368 GTGCGGTACGCTGCTGTTGACGAC---AACACACGCGCGGTTGGACGATCTCTTG 424
Db 129 ValIleValArgLeuLeuLeuAspAspPheAsnSerSerProGluLeuAspGlnLeuLeu 148
QY 425 CTCGCGCTCGACAGCAGCTCCCAATATCGAAGTGGCGCTTTCACACCCCTCTCTCTACGC 484
Db 149 LeuArgIleSerLysHisLysAsnIleAlaValArgLeuIleAsnProMetProTyrArg 168
QY 485 AAATGGCGGCACTCGGTACTCGACCGACTTCCCGCGCTCAACCGCGCATGCACAAAC 544
Db 169 GlyPheArgSerLeuAsnTyrMetLeuHisProIleArgThrAsnArgMetHisAsn 188
QY 545 AAATCTTTTACCGCGACACCGCGCCACCATCTCGGCGGAGCGCAATATCGCGACGAA 604
Db 189 LysSerMetThrPheAspAsnLysIleSerValIleGlyArgAsnIleGlyAsnGlu 208
QY 605 TACTTCAAAGTCGGTGAGACACCGTTTTCGCGACCTGGACATCTCTGCCACCGCGACG 664
Db 209 TyrLeuAsnAsnAlaProValAsnHisPheAlaAspLeuAspValMetLeuValGlyHis 228
QY 665 GTCTGCGGGAAGTATCGCAGCTTCGACCGCTACTGGCAAGCATTCGCCCCACAAAC 724
Db 229 ValValGlyLysIleThrGlnSerPheGluIleTyrTrpAlaSerProLeuSerPheAsp 248
QY 725 GCCACGCGCATCTCGGACGGGCAACATCGGC-----AAGGTCCTCAAGCAGCTCGGA 778
Db 249 IleGluThrLeuValLysHisAspAsnLysAspIleSerGlyIleLysProValVal 268
QY 779 TACAACGAC-----GAAACATCCAGACACGCGCTCTCTCGGCTAC 817
Db 269 PheAspGluLeuGluLysValLysAspSerSerAspAlaAspArgGluLeuArgThrTyr 288
QY 818 CCGCAAAACCGTCGACAGTCGCCCTCTACCAAAAATAACAGACGGGACGATCGAGTGG 877
Db 289 ArgGlnAlaMetGlnAsnSerThrIleGlyGlnAspLeuAlaGlnGlnValProPhe 308
QY 878 CAGAGGTCAAACCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGC 937
Db 309 PheTrpThrGluIleAspLeuIleAlaAspAsnValIleLysLeuAsnGlyHisSerAsp 328
QY 938 CGCAAAACCGCATTCGCGGAGGTGCAAGACGGCTCAAAACAGCCCCGAAAAGCGTC 997
Db 329 ProAsnGluPheLeuValSerGlnLeuGlnThrGlnLeuGlnProSerLysLysLeu 348
QY 998 TATCTGTTTCCCTTATTTCTGCTTACAAAATCCGGCACAGACGCACTGCGCAAACTG 1057
Db 349 SerIleIleSerTyrPheValProThrLysGluGlyIleAspThrLeuIleThrLeu 368
QY 1058 GTGACGAGCGCATAGCGTTTACCGTCTCGACCAACTCGCTACAGCGGACCGAGCTTGC 1117
```

```
Db 369 AlaLysMetGlyValGluValArgIleLeuThrAsnSerPheAspAlaThrAspValGly 388
QY 1118 GCCCTCATTCGGCTACGTCAAAATACCGAAAAACCGTGTCTCAAGCCGGCATCAAACTC 1177
Db 389 IleValHisAlaGlyTyrAlaHisTrpArgLysGlnLeuLeuAlaAlaGlyValHisLeu 408
QY 1178 TACGAGCTGCAACCCCAAC---CATGCCGTCCCGCCACAAAAGACAAA-----GGCCTG 1228
Db 409 PheGluIleLysSerSerAlaGlnSerIleGlnAspAsnGluAsnArgPheTrpArgThr 428
QY 1229 ACCGCGACCTCGTAACACGAGCTGCATGCCAAAACCTTCATTGTGCGCGCAACGCAATC 1288
Db 429 ArgGlnHisSerThrThrSerLeuHisAlaLysAlaPheAlaValAspAspGlnIle 448
QY 1289 TTCATCGCTCATTCACACCTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCGTC 1348
Db 449 PheIleGlySerTyrAsnValAspProArgSerAlaAsnIleAsnThrGluLeuGlyVal 468
QY 1349 GTCATCGAAAGCCCAAAATCGAGACAGATGGAGCGCACCTCGCGCATACC-----1402
Db 469 LeuIleLysAspSerLysLeuAlaGlyGlnLeuHisLysAlaLeuSerAsnSerGlnAla 488
QY 1403 ACACCGGAATACGCTACCGCTTACCCTCGACAAACACAAACCGCTCGCAATGGCAGAT 1462
Db 489 IleThrHisGlnAlaTyrGluLeuLysLeuAspAlaLysGlyAsnIleAsnTrpHis--- 507
QY 1463 CCGCGCACCCGAAAAAACCTACCCGAAAC-----GAACCCGAA 1498
Db 508 -----ThrIleGluAsnGlyGlnProValIleLeuHisGluProHis 522
QY 1499 GCCAAACTTTGAAACGATCGCGCAAAAATCCTATCCCTGCTGCCCATCGAAGTTTA 1558
Db 523 MetAsnHisLysAspArgIleIleTrpLeuAlaGlyLeuMetProIleAspTrpLeu 542
QY 1559 TTA 1561
Db 543 Leu 543
RESULT 7
ABO75163
ID ABO75163 standard; protein; 564 AA.
XX ABO75163;
AC ABO75163;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #7338.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 98US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI: 2003-615309/58.
DR N-PSDB; ABD08734.
XX
Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
```



```

XX AC AAW55452;
XX DT 24-JUN-1998 (first entry)
XX DE H. pylori ORF 02aell612_33203250_cl_51 secreted protein.
XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
XX KW identification; binding compound; bacteria; life cycle; activator;
XX KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
XX OS Helicobacter pylori.
XX PN WO9737044-Al.
XX PD 09-OCT-1997.
XX PF 27-MAR-1997; 97WO-US005223.
XX PR 29-MAR-1996; 96US-00625811.
XX PR 02-APR-1996; 96US-00758731.
XX PR 25-OCT-1996; 96US-00736905.
XX PR 28-OCT-1996; 96US-00738859.
XX PR 06-DEC-1996; 96US-00761318.
XX PA (ASTR ) ASTRA AB.
XX PI Smith D, Alm RA;
XX DR WPI; 1997-503122/46.
XX DR N-PSDB; AAV24861.
XX PT Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) -
XX PT useful in vaccines to treat or prevent H. pylori infection and for
XX PT diagnosis of H. pylori infection.
XX PS Claim 14,94; Page 659-660; 1145pp; English.
XX CC This sequence is a H. pylori secreted protein. The protein may be used in
XX CC a vaccine to prevent or treat H. pylori infection or to identify H.
XX CC pylori polypeptide binding compounds, useful as potential H. pylori life
XX CC cycle activators or inhibitors. The DNA and probes derived from it may be
XX CC used for the identification of H. pylori in a sample and the diagnosis of
XX CC H. pylori infection. Nucleic acid sequences complementary to the DNA act
XX CC as antisense sequences and can be used to prevent the translation of H.
XX CC pylori mRNA. Antibodies against the protein can be used in immunoassays
XX CC to evaluate the abundance and distribution of H. pylori-specific
XX CC antigens. The genomic sequence of H. pylori (ATCC 55679) was determined
XX CC from overlapping contigs generated by mechanically shearing the bacterial
XX CC DNA. The sequences were analysed for ORF of at least 180 nucleotides, and
XX CC the predicted coding regions defined by computer evaluation. To identify
XX CC likely H. pylori antigens for vaccine development, the amino acid
XX CC sequences predicted from various ORF were analysed for significant
XX CC homology to other known or exported membrane proteins. Having identified
XX CC and determined the sequences of interest, particular regions can be
XX CC isolated from H. pylori by PCR amplification for recombinant polypeptide
XX CC production, e.g. in E. coli hosts
XX SQ Sequence 502 AA;

Alignment Scores:
Pred. No.: 7.88e-38 Length: 502
Score: 632.00 Matches: 161
Percent Similarity: 51.7% Conservative: 93
Best Local Similarity: 32.8% Mismatches: 181
Query Match: 22.2% Indels: 56
DB: 2 Gaps: 17

US-10-665-990A-13 (1-1561) x AAW55452 (1-502)
QY 176 ATCCGGCACCCCTCATCAACACGGGCTATCCGACATCTAC----- 217
Db 30 IleSerTyrAspProTyrThrThrThrIleGlySerLeuTyrAlaLysAsnLeuLysGlu 49

```

```

QY 218 -----CTGCTCAGCAGACCCCGCAGAGCCCTTGGCCCGCCG 253
Db 50 AsnProLysHisSerAlaAlaIleLeuLeuGluAspGlyPheAspAlaLeuHisArg 69
QY 254 GCCGCCCTTATCGAATCTGCCGAACACAGCCTCGATTGGCAATACATACATTTGGCGCAAC 313
Db 70 ValGlyLeuIleArgMetSerGlnLysSerIleAspMetGlnThrTyrIleTyrLysAsn 89
QY 314 GACATTTCCGGCAGGCTCTGTTCACCTCATGTACCTTGGCGCAGAGACGGCGCTGGCG 373
Db 90 AspLeuSerSerGlnValIleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLys 109
QY 374 GTACGCCTGTCTGGACGACAAACACCGCGCGGTTGGAC-----GATCTC 421
Db 110 ValArgIleLeuLeuAspAsn-----GlyLeuAspSerAspPheSerAspIle 126
QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGCCTGTTTCAACCCCTTCGTCCTA 481
Db 127 MetLeu-----LeuAsnPheHisLysAsnIleGluValLysIlePheAsnProTyrTyrIle 145
QY 482 CGCAAAATGGCGCAGCTCGGCTAC-----CTGACCGACTTCCCGCGCTCAACCGC 532
Db 146 ArgAsn---LysGlyLeuArgTyrPheGluMetLeuAlaAspTyrGluArgIleLysLys 164
QY 533 CGCATGCACAACAATCTTTACCGCCGACACCGCGCCACCATACTCGGCGGAGCGCAAT 592
Db 165 ArgMetHisAsnLysLeuPheIleValAspAsnPheAlaValIleIleGlyGlyArgAsn 184
QY 593 ATCGCGCAGCAATACTTCAAAAGTCGGTGAGACACCGTTTTCGCGCAGCTCGACATCTCTC 652
Db 185 IleGlyAspAsnTyrPheAspAsnAspLeuAspThrAsnPheLeuAspAspAlaLeu 204
QY 653 GCCACCGCAGCGTCGTGCGGCAAGTATCGCAGCACTTCGACCGCTACTGGCAAGCCAT 712
Db 205 PhePheGlyGlyValAlaSerLysAlaLysGluSerPheGluAsnTyrTrpArgPheHis 224
QY 713 TCCGCCACACAGCCCGCATCATCTCGCAGC-----GGC 748
Db 225 ArgSerIleProValSer---LeuLeuArgThrHisLysArgLeuLysAsnValLys 243
QY 749 AACATCGCGCAGGGCTTTCAGACACTCGGATACACAGCAACATCATCAGACACGCGCTC 808
Db 244 GluIleAlaLysLeuHisGlyLysIleIleSerIleAlaGluAspAlaAsnGluPheGlu 263
QY 809 CTGCGCTACCGCGAAACCGTCGACAGTCGCGCCCTTACCAA---AAATACAGACGGGA 865
Db 264 LysLysValAsnAspPheIleGluArgPheGlnLysTyrGlnTyrProIleTyrTyrGly 283
QY 866 CGCATCGACTGGCAGAGGTCCAAACCGCTGATCAGGACACCCCTGCAAAAGACTC 925
Db 284 -----AsnAlaIlePheLeuAlaAspLeuProAlaLys---Ile 295
QY 926 GACCGCGACCGCGCAACCGCGGAGTTCGCGGAGGTGCAAGACGCGCTCAAAACAGCCC 985
Db 296 AspThrProLeuTyrSer---ProIleLysIleAlaPheGluLysAlaLeuLysAsnAla 314
QY 986 GAAAAAAGCGTCTATCTGGTTTCACCCCTATTTTCGTCCTTACAAATCCGGCAGACGCA 1045
Db 315 LysAspSerValPheIleAlaSerSerTyrPheIleProGlyLysLysIleMetLysIle 334
QY 1046 CTGGCAAAATCGTCGAGGACGGCATAGACGTTCAGTCTTCCACCACTCGCTACAGGCG 1105
Db 335 PheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeuThrAsnSerLeuSerSer 354
QY 1106 ACCGACGTTCGCGCGCTCCATTCCCGCTACGTCAATACCGAAACCGCTGTCTCAAAGCC 1165
Db 355 ThrAspAlaIleValValTyrGlyAlaTrpGluAArgTyrArgAsnLysLeuValArgMet 374
QY 1166 GGCATCAAACTCTAGAGTGCACCCCAACCATGCGTCCCGCCGACCAAAACCAAGC 1225
Db 375 GlyAlaAsnValTyrGluIleArgAsnAspPhePheAsnArgGlnIleLysGlyArg--- 393

```



```
Db      264  LysLysValAsnAspPheIleGluArgPheGlnLysTyrGlnTyrProIleTyrTyrGly 283
QY      866  CGCATCGACTGCGAGAGCGTCAAAACCGCCTGATCAGCACACCCCTGCAAAAGGACTC 925
Db      284  -----AsnAlaIlePheLeuAlaAspLeuProAlaLys---Ile 295
QY      926  GACCGCGACCGCGGAAACCGCGATTCGGGAGGCTGCAAGACGCGCTCAAAAGGCC 985
Db      296  AspThrProLeuTyrSer---ProIleLysIleAlaPheGluLysAlaLeuLysAsnAla 314
QY      986  GAAAAAGCGTCTACTGTTTCAACCTATTTCGCTTACAAATCCGGACACAGCCA 1045
Db      315  LysAspSerValPheIleAlaSerSerTyrPheIleProGlyLysLysIleMetLysIle 334
QY      1046  CTGGCAAAACTGGTCAGACGGCATAGACGTTACGTCCTGACCAACTCGCTACAGGCG 1105
Db      335  PheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeuThrAsnSerLysSer 354
QY      1106  ACCGACGTTGCGCGCTCATTCGGGTACGTCAAAATACCGAAACCGCTGCTCAAGCC 1165
Db      355  ThrAspAlaIleValValTyrGlyAlaTrpGluArgTyrArgAsnLysLeuValArgMet 374
QY      1166  GGCATCAAACTCTACGAGCTGCAACCCCAACCATCGCTCCCGCCACAAAGACAAAGGC 1225
Db      375  GlyAlaAsnValTyrGluIleArgAsnAspPheAsnArgGlnIleLysGlyArg--- 393
QY      1226  CTGACCGGCGAGCTCGGTAAACCAAGCTGATGCGCAAAACCTTCATGTCGACGCAACGC 1285
Db      394  -----PheSerThrLysHisSerLeuHisGlyLysThrIleValPheAspAlaLeu 411
QY      1286  ATCTTCATCGGCTCATTTCAACCTCGACCCCGCTTCGCGACGCTCAATACCGAAATGGGC 1345
Db      412  ThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyrIleAsnThrGluSerAla 431
QY      1346  GTCGTCATCGAAGCCCAAAATCCGACAAAGATGGAGCGCACCTCGCCATACCACA 1405
Db      432  ValLeuPheAspAsnProSerPheAlaLysArgValArgLeuSerLeuLysAspHisAla 451
QY      1406  CCGCAATACGCTACCGCTTACCCTCGACAAACACACACCGCTGCAATGGCAGCATCCC 1465
Db      452  ---GlnGlnSerTrpHisLeuValLeuTyrArgHis---ArgValIleTrpGlu----- 467
QY      1466  GCCACCGGAAACCTTAC-----CCGAAGAACCGAAGCCAACTTTGGAAA 1513
Db      468  AlaThrGluGluGlyIleLeuIleHisGluLysAsnSerProAspThrSerPhePheLeu 487
QY      1514  CGCATCGCGCGCAAAATCCTATCCCTGCTGCC 1546
Db      488  ArgLeuIleLysGluTrpSerLysValLeuPro 498

RESULT 10
ID AAW98378 standard; protein; 428 AA.
XX
AC AAW98378;
XX
XX
DT 31-MAR-1999 (first entry)
XX
DE H. pylori GHPO 1375 protein.
XX
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease.
XX
OS Helicobacter pylori.
XX
PN WO9843478-Al.
XX
PD 08-OCT-1998.
XX
PF 01-APR-1998; 98WO-US006371.
XX
XX 01-APR-1997; 97US-00833457.
PR
```

```
PR      24-JUN-1997; 97US-00881227.
PR      29-JUL-1997; 97US-00902615.
XX
XX (INNR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
XX WPI; 1998-542293/46.
XX N-PSDB; AAX14097.
PT New isolated Helicobacter polynucleotides - used to develop products for
PT the diagnosis, prevention and treatment of Helicobacter infections and
PT gastrointestinal diseases.
XX
XX Claim 8; Page 581-583; 2054pp; English.
XX
XX This sequence represents a Helicobacter pylori GHPO protein of the
XX invention. The polypeptides can be used for preventing or treating
XX Helicobacter infections, and gastroduodenal diseases associated with
XX these infections, including acute, chronic, and atrophic gastritis, and
XX peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
XX used for the production of antibodies. The products can also be used for
XX detection and diagnosis
XX
XX SQ Sequence 428 AA;

Alignment Scores:
Pred. No.: 9,4e-35 Length: 428
Score: 590.00 Matches: 145
Percent Similarity: 50.1% Conservative: 80
Best Local Similarity: 32.3% Mismatches: 146
Query Match: 20.7% Indels: 78
DB: 2 Gaps: 14

US-10-665-990A-13 (1-1561) x AAW98378 (1-428)
QY      272  GCCGAACACAGCCTCGATTTCGAATCTACATTTGGCGCAACGACATTTCCGGCAGGCTG 331
Db      2 SerGlnLysSerIleAspMetGlnThrTyrIleTyrLysAsnAspLeuSerGlnVal 21
QY      332  CTGTTCAACCTCATGTACTCTTCCCGCAAGACGCGCGTGGCGTACGCTGCTGTTGGAC 391
Db      22 IleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLysValArgIleLeuLeuAsp 41
QY      392  GACAAACAACGCGCGGTTGGAC-----GATCTCTCTGCTCGCCCTCGACAGC 439
Db      42 AspAsn-----GlyLeuAspSerAspPheSerAspIleMetLeu---LeuAsnPhe 57
QY      440  CATCCCAATATCGAAGTCGGCTGTTCAACCCCTTCGTCTCTACGCAAAATGGCGCGACTC 499
Db      58 HisLysAsnIleGluValLysIlePheAsnProTyrTyrIleArgAsn---LysGlyLeu 76
QY      500  GGCTAC-----CTGACCGACTTCCCGCCCTCAACCGCGCATGCACACAAATCC 550
Db      77 ArgTyrPheGluMetLeuAlaAspTyrGluArgIleLysLysArgMetHisAsnLysLeu 96
QY      551  TTTACCGCGCAACACCGCGCCACCATCTCTCGCGGACGCAATATCGGCGACCAATCTTTC 610
Db      97 PheIleValAspAsnPheAlaValIleIleGlyGlyArgAsnIleGlyAspAsnTyrPhe 116
QY      611  AAGTCGGTGAGGACACCGTTTTCGCCGACCTGGACATCCTCGCCACCGCAGCGCTGTC 670
Db      117  AspAsnAspLeuAspThrAsnPheLeuAspAlaLeuPhePheGlyGlyValAla 136
QY      671  GCGGAAGTATCGCAGCACTTCGACCGCTACTGGGCAAGCCATTCGCCCCACCAACCCAGC 730
Db      137  SerLysAlaLysGluSerPheGluArgTyrTrpArgPheHisArgSerIleProValSer 156
QY      731  CGCATCTCCGCGAGCGGCAACATCGGCAAGGGTCTTCAAGCACTCGGATACACGACGAA 790
Db      157  ---LeuLeuArgThr----- 160
```


QY 524 CTCACCGCGCATGCACAAACAAATCTTTTACCGCGCAACACCGCGCCACCCATCTCGGC 583
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
191 IleAsnArgArgMetHisAsnLysSerPheIleAlaAspAsnGlnValAlaLeuIleGly 210
QY 584 GGACGCAATATCGCGCAGCAATCTTCAAAGTCGGTGAGGACACCGTTTTCGCCACCTG 643
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
211 GlyArgAsnMetThrAsnGlnTyrTyrAsnValSerAspSerTyrGlnPheSerAspVal 230
QY 644 GACATCTCTGCCACCGCAGCGTCGTCGGCGAAGTATCGCAGCACTTCGACCGCTACTGG 703
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
231 AspValMetLeuValGlyAlaAlaValAspAspIleValAsnSerPheValAspTyrTrp 250
QY 704 GCAAGCATTCTGCCACACGCCACGCCGCGCATCATCGC 742
Db 251 AsnHisGluTyrAlaTyrSerValGlnSerIleValArg 263
RESULT 12
ADA34304
ID ADA34304 standard; protein; 275 AA.
XX
AC ADA34304;
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #1465.
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR N-PSDB; ADA30178.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 5591; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 275 AA;

Alignment Scores:
Pred. No.: 1.38e-26 Length: 275
Score: 478.50 Matches: 100
Percent Similarity: 61.9% Conservative: 51
Best Local Similarity: 41.0% Mismatches: 78
Query Match: 16.8% Indels: 15
DB: 6 Gaps: 5

US-10-665-990A-13 (1-1561) x ADA34304 (1-275)
QY 866 CGCATCGACTGGCAGAGCGTCCAAACCCGCTGTACGCGACACCCCTGCAAAAGAGCTC 925
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
35 LysPheAspTrp-----ValLysAlaGluValValLysAspSerProAspLysIleArg 52
QY 926 GACCGCAGCCGCGCAACCCCGGATTCCCGGGAGGTCGCAAGACGCGCTCAAAAGCC 985
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
53 SerLysAlaLysLysGluGluHisLeuAsnPheGlnLeuIleAsnHisLeuGluLysPro 72
QY 986 GAAAAACCGCTATCTCTGTTTCACCTTATTCGTCCTACAAAATCCGGACACGACACGCA 1045
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
73 GluSerAsnValAspLeuIleSerAlaTyrPheIleProGluLysGlnGlyAlaLysIle 92
QY 1046 CTGGCAAAACCTGGTCAGGACGCGATACGCTTACCGTCTCCTACCAACTCGCTACAGCG 1105
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
93 LeuSerThrLeuAlaLysGluGlyValGluValArgValLeuThrAsnSerPheLysAla 112
QY 1106 ACCGACGTTGCGCGCTCCATTCCGGCTACGTCAAATACCGAAAACCGCTGCTCAAAGCC 1165
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
113 AsnAspValAlaValHisAlaPheTyrGlyLysTyrArgLysGluLeuLysAsn 132
QY 1166 GGCATCAAACTCTAGGACTGCAACCC-----AACCAT 1198
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
133 GlyValGlnLeuTyrGluPheLeuProThrProAspLysArgAspLeuAsnLysAsnThr 152
QY 1199 CGCGTCCCGCCACAAAAGACAAA---GGCCTGACCGCGCAGCTCCGTAACCGCCTGCAT 1255
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
153 AspGluLeuAlaThrLysAlaLysValAsnMetLysGlyLeuSerArgSerSerLeuHis 172
QY 1256 GCCAAAACCTTCATTGTGGACGGCAACGCACTTCATCGGCTCATCACTCGACGCC 1315
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
173 ThrLysLeuMetAlaLeuAsp---GluGlnValPheIleGlySerPheAsnPheAspPro 191
QY 1316 CGTTCGCGACGCTCAATACCGAAATGGCGTCGTCATCGAAAGCCCAAAATCCAGAA 1375
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
192 ArgSerAlaTyrLeuAsnThrGluIleGlyValIleLeuAspSerProSerLeuAlaLys 211
QY 1376 CAGATGAGCGCACCCTCGCGCATACACACCGCAATACGCTACCGGTTTACCTCGAC 1435
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
212 ThrIleHisHisThrMetAspGluAsnLysTyrAlaTyrLysLysLeuAsp 231
QY 1436 AAACACAACCGCTGCAATGGCAGAT-----CCGCGCACCCGAAAAACCTACCCGAC 1489
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
232 ProAsnAsnHisIleTyrTrpGlnGlnGluThrProLysGlyProValIleTyrLysLys 251
QY 1490 GAAACCGGAAGCCAAACTTTGGAAACGCGATCGCGCAAAATCCTATCCTGCTGCCCATC 1549
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
252 GluProGluMetLysTrpTrpGlnLysAlaGlyMetLysLeuLeuSerTrpLeuProLeu 271
QY 1550 GAAGGTTTATTATTA 1561
Db :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
272 GluGlyPheMet 275
RESULT 13
ABM88218
ID ABM88218 standard; protein; 2519 AA.
XX
AC ABM88218;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice abiotic stress responsive polypeptide SEQ ID NO:6464.
XX
KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX


```
QY 1367 ATCGAGAACACATGAGCGCACCTCGCCGATACACACCCGAATACGCCTACCGCGTT 1426
Db |||||:::
837 PheAlaGluGlnLeu-----AlaGluSerIle 845
QY 1427 ACCCTCGACAAACAAACCGCTGCAATGCGACGATCCGCGCACCCGAAACCT 1481
Db |||||:::
846 LeuArgAspIleHisProGlnAsnSerTrp-ThrValAlaProArgAlaLysPro 863

RESULT 14
ADA36398
ID ADA36398 standard; protein; 321 AA.
XX
AC ADA36398;
XX
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #3559.
XX
XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
XX Acinetobacter baumannii.
XX
XX US6562958-B1.
XX
XX 13-MAY-2003.
XX
XX 04-JUN-1999; 99US-00328352.
XX
XX 09-JUN-1998; 98US-0088701P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton G, Bush D;
XX
XX WPI; 2003-576092/54.
XX
XX N-PSDB; ADA32272.
XX
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
XX Example; SEQ ID NO 7685; 328pp; English.
XX
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
XX
SQ Sequence 321 AA;

Alignment Scores:
Pred. No.: 1.88e-21 Length: 321
Score: 409.00 Matches: 100
Percent Similarity: 50.0% Conservative: 44
Best Local Similarity: 34.7% Mismatches: 110
Query Match: 14.3% Indels: 34
DB: 6 Gaps: 6

US-10-665-990A-13 (1-1561) x ADA36398 (1-321)

QY 74 CTTCTCTGTTCATGTTCTTCATGGTTG-----CCCCCACTGGAAGACGGAGGAACG 127
Db :::: |||||
44 IleValIleLysThrAsnAsnTrpLeuAsnAspGluAlaAlaGluGlnLysThrGln--- 62
QY 128 CGTCATTCAATACTTCCAAACCTGTCTCTCTGGACACATCTCTGCAATTCGGCACACC 187
```

```
Db 62 -----
QY 188 CQTATAACAACGGCTATCCGACATCTACCTGCTCGACGACCCCCACGAAAGCCCTTGCC 247
Db |||||:::
63 -----GlnGlyLeuThrAlaPheValAlaLeuAspAlaPheMetSerIleAla 79
QY 248 GCCCGCGCGCCCTTATCGAATCTGCCGAACACAGCCTCGATTTCGAATACTACTATTTGG 307
Db |||||:::
80 SerArgIleTyrLeuIleArgAsnAlaLysGluThrIleAspLeuGlnTyrTyrIleTrp 99
QY 308 CGCAACGACATTTCCGCGACGCTGCTTCAACCTCATGTACCTTGCCTGCGCAAGCCGCG 367
Db |||||:::
100 ThrAsnAspPheValGlyAsnLeuIleHisGluLeuLeuLysAlaAlaAspArgGly 119
QY 368 GTGCGGTACGCTGCTGTGGACGACAAACACGCGCGGGTTGGACGATCTCTCTGCTC 427
Db |||||:::
120 IleLysValArgLeuLeuIleAspAspGlnAsnGlyIleLysLeuAspGlyIleLeuArg 139
QY 428 GCCCTCGACAGCCATCCCAATATCGAAGTGGCCTGTTCAACCCCTTCGTCTCTAGCAA 487
Db |||||:::
140 SerLeuLeuGlnHisThrAsnPheGluIleArgLeuPheAsnProTyrLysPheArgTyr 159
QY 488 TGGCGCGCACTCGGCTACCTGACCGACTTCCCGCCCTCAACCGCGCATGCACAAACAA 547
Db |||||:::
160 LeuArgIlePheAspTyrLeuPheArgPheLysLysValAsnHisArgMethHisAsnLys 179
QY 548 TCCTTTACCGCGCAACCGCCACCATATCTCGCGGACGCAATATATCGGCGACGAATAC 607
Db |||||:::
180 LeuIleIleAlaAspAlaSerIleAlaValThrGlyGlyArgAsnIleSerSerGluTyr 199
QY 608 TTCAAAGTCGGTAGGACACCGTTTTCGCGACCTGGACATCTCGCCACCGCGACGCTC 667
Db |||||:::
200 PheGluAlaSerSerLysPheGlnPheThrAspMetAspIleLeuPheTyrGlyHisAla 219
QY 668 GTGCGCGAAGTATCGCACGACTTCGACCGCTACTGGCAAGCATTCGCGCCCAACACGCC 727
Db |||||:::
220 ValArgHisAlaGlnAlaValPheThrAspPheTrpGluSerThrLeuSerValAsnAla 239
QY 728 AGCGCATCATCCGCGAGCGCAACATCGGCAAG---GGTCTTCAAGCACTCGGATACAAAC 784
Db |||||:::
240 ThrGluIleIle-----GlyThrCysAlaGluHisHisLeuLysAlaLeuArgGluHis 257
QY 785 GACGAAACA-----TCCAGACACGCGCTCTCGCTACCGCGAAACCGTCGAAACAG 835
Db |||||:::
258 TyrGluGlnLeuHisHisGluAspHisSerLeuThrGluAspLysLeuTyrAspAlaGln 277
QY 836 TCGCCCTCTCTACCAAAAAATACAGAGCGGACGCGCATCGATCGAGAGCGTCCAAACCCGC 895
Db |||||:::
278 SerTyrLeuLysGluLeuLeuGluHisAsnProIleGlnTrpSer-----LysAlaHis 295
QY 896 CTGATCAGCGACACCCCTGCAAAA 919
Db |||||:::
296 PheValAlaAspSerProLysLys 303

RESULT 15
AEA04533
ID AEA04533 standard; protein; 5178 AA.
XX
XX AEA04533;
XX
DT 28-JUL-2005 (first entry)
XX
DE Human protein from gene under-expressed in cancer, MUC2.
XX
XX Tumor marker; colon tumor; cancer; cytostatic; neoplasm; diagnostic;
XX microarray; drug screening.
XX
XX Homo sapiens.
XX
XX WO2005044990-A2.
XX
XX 19-MAY-2005.
```


XX 01-NOV-2004; 2004WO-US036404.
 XX PF

XX 04-NOV-2003; 2003US-00700439.
 XX PR

XX (FARB) BAYER HEALTHCARE LLC.
 XX PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

XX Burgess C, Myerow S, Thiagalingam A, Maimonis P, Molino G;
 XX PI Burgett L, Boardman LA, Thibodeau S, Lewis M;

XX WPI; 2005-372198/38.
 XX DR N-PSDB; AEA04440.

XX DR REFSEQ; NP_002448.
 XX DR

PT Detecting expression of one or more nucleic acid sequences in biological
 PT sample, useful for detecting cancer, comprises detecting a change in the
 PT expression level of one or more nucleic acid sequences relative to a
 PT control expression level.
 XX

PS Claim 20; SEQ ID NO 178; 256pp; English.
 XX

CC The invention relates to detecting differential expression of one or more
 CC nucleic acid sequences (appearing as AEA04356-AEA04448 in a biological
 CC sample comprising obtaining the sample from a subject, and detecting a
 CC change in the expression level of one or more nucleic acid sequences
 CC relative to a control expression level of the nucleic acid sequences
 CC new. Also included are detecting cancer (or a pre-malignant condition
 CC thereof) in a subject (comprising comparing the expression level of one
 CC or more nucleic acid sequences in a biological sample from the subject
 CC with a control expression level of the nucleic acid sequences, where a
 CC change of at least two-fold in the expression level of the nucleic acid
 CC sequences is indicative of cancer or pre-malignant condition), monitoring
 CC the onset (or progression, or regression) of cancer (or a pre-malignant
 CC condition) in a subject (by detecting in a biological sample of the
 CC subject at a first point in time the expression of one or more nucleic
 CC acid sequences, repeating the first step at a subsequent point in time
 CC and comparing the expression level detected, where a change in the
 CC expression level is indicative of progression of cancer or its pre-
 CC malignant condition in the subject), determining prognosis for cancer or
 CC its pre-malignant condition in a subject (comprising detecting in a
 CC biological sample of the subject, the expression level of one or more
 CC nucleic acid sequences, comparing the expression level with a reference
 CC expression level of the nucleic acid sequences and evaluating the
 CC prognosis of the subject based on the comparison), determining the
 CC efficacy of a test compound for inhibiting cancer in a subject,
 CC determining the efficacy of a therapy for inhibiting cancer in a subject,
 CC selecting a composition for inhibiting cancer in a subject, inhibiting
 CC cancer in a subject, a polypeptide encoded by the nucleic acids above
 CC (appearing as AEA04449-AEA04541), an antibody that specifically binds to
 CC the polypeptide sequence, and detecting in a biological sample the
 CC presence of a polypeptide. The method is useful for detecting in a
 CC differential expression of one or more nucleic acid sequences in a
 CC biological sample, which is useful for detecting cancer (especially colon
 CC cancer), monitoring the onset, progression, or regression of cancer or a
 CC pre-malignant condition, or determining prognosis for cancer or its pre-
 CC malignant condition in a subject, or for determining the efficacy of a
 CC test compound for inhibiting cancer in a subject. The compound is useful
 CC for inhibiting cancer in a subject. The antibodies may also be used to
 CC treat cancer. The present sequence is a protein from a human gene under-
 CC expressed in cancer samples.
 XX

XX Sequence 5178 AA;

Alignment Scores:

Pred. No.:	8.08e-16	Length:	5178
Score:	336.50	Matches:	141
Percent Similarity:	33.6%	Conservative:	55
Best Local Similarity:	24.1%	Mismatches:	198
Query Match:	11.8%	Indels:	190
DB:	9	Gaps:	22

US-10-665-990A-13 (1-1561) x AEA04533 (1-5178)

QY	119	ACGGAAGCGCTCATTTCAAT-----ACTTCCAAACCTGTCTCTCGACAC	166
Db	1248	ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh	1267
QY	167	ATCTGCAAAATCCGGCACACCCCTCATAAACAACGGGTATCCGACATCTACTGCTCGAC	226
Db	1267	rPheThrThrIleThrLeuThrThrProThrProThrSerPheThrThrThrThrTh	1287
QY	227	GACCCCAACGAGCCCTTGGCCGCCGCCCTTATCGAATCTGCCGAACACAGCCTC	286
Db	1287	rThrProThr-----SerSerTh	1293
QY	287	GATTGTGAATATACATTTTGGCGCAACGACATTTCCGGCAGCGCTGCTTCAACCTCA--	344
Db	1293	rValLeuSerThrThr-----ProLysLeuCysCysLeuThrProSerAs	1307
QY	344	-----	344
Db	1307	pTIpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs	1327
QY	345	-----TGTACCTTGGCGCAGAACGGCGGTGGCGGTACGCC-----	380
Db	1327	pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe	1347
QY	380	-----	380
Db	1347	rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs	1367
QY	380	-----	380
Db	1367	nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs	1387
QY	381	TGCTGTGG-----ACGACAACAAC	400
Db	1387	nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrProSe	1407
QY	401	ACGGCGGGTGGACGATCTCTGCTCGCCTCGACAGCCATCCCAATATCGAAGTGGC	460
Db	1407	rProProProThrThrThrThrLeuProProThrThrPro-----	1422
QY	461	CTGTTCACCCCTTCGTCTACGCAATGGCGGCACCTCGGCTACGACGACTTCCCC	520
Db	1423	-----SerProProThrThrThrThrProProProThrThrThrProSerProPr	1441
QY	521	CGCTCAACCGCGCATGCACAACAAATCCTTTACCGCCGACACACCGCCGCCCATCTC	580
Db	1441	oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro-----	1456
QY	581	GGCGGACGCAATATCGCGCAGCAATCTTCAAAGTCGTGAGGACACCGTTTCGCCGAC	640
Db	1457	-----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh	1474
QY	641	CTGGACATCTCGCCACCGCAGCGTGTGGCGGAGTATCGCACGACTTCGAC-----	695
Db	1474	rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr	1494
QY	696	-----GCTACTGGGCAACCATTCGCCCAACAAACGCCGCGCATCTCCGCGCGGC	748
Db	1494	oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh	1514
QY	749	AACATCGGCAAGGTCTTCAAGCAGCTCGGATACAAAGCAACAAACATCCAGACACGCGTC	808
Db	1514	rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrThrTh	1529
QY	809	CTGCGCTACCGGAAACCGTCGAAC-----AGTCGCCCTCTACCAAAAAATACAGACG	862
Db	1529	rThrThrThrProProProThrThrThrProSerProProThrThr-----	1544
QY	863	GGAGCATCGACTGGCGAGCGTCCAAACCCGCTGATCAGCGACACCCCTCGAAAAAGGA	922
Db	1545	-----ThrProIleThrProPr	1550

```
QY 923 CTCGACCGCGCGGCGGAAACCGCGGATTGCGGGAGGCTGCAAGACGCGCTCAAAACAG 982
Db : : : : : |||
1550 oThrSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1563
QY CCGGAAAAAAGCGCTCTATCTGTTTCACCTATTTCGTCCTTACAAAATCCGGCACAGAC 1042
Db : : : : : |||
1563 rProProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrTh 1583
QY GCACGTGGCAAAACTGGTCAGACGCGCATAGAGTTACCGTCTCGACCA-----AC 1093
Db : : : : : |||
1583 r-----ThrThrProSerProProThrThrThrThrThrThrThrThrThrThr 1593
QY TCGCTACAGCGACGCGAGTTGCCCGCGCTCCATTCGGGTAGTCAAAAT----- 1142
Db : : : : : |||
1593 rThrThrProProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrPr 1613
QY : : : : : |||
1143 -----ACGAAAAACCGCTGCTCAAAAGCGGCATCAAACTC 1177
Db : : : : : |||
1613 oProThrThrThrProSerProProThrThrThrThrThrThrThrThrThrThrThrTh 1633
QY TACGAGCTGCAACCAACGATCGCTCCCGCCCAAAAAGACAAAGAGGCTGACCGGACG 1237
Db : : : : : |||
1633 rThrLeuProProThrThrThrProSerProProThrThr-----ThrThrTh 1650
QY TCGGTAAACGAGCTGCATGCCAAACCTTCATTGTGGACGGCAACGCGATCTTCATCGCG 1297
Db : : : : : |||
1650 rProProProThrThrThrProSerProProThr---ThrThrThrProSerProProIl 1669
QY TCATTCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGTTCGTATCGAA 1357
Db : : : : : |||
1669 eThrThrThrThrThrThrProProProThrThrThrThrProSerProThrThrThrPr 1689
QY AGCCCCAAATCGAGACAGATGAGCGCACCTCGCGGATACACACCCGGAATACGCC 1417
Db : : : : : |||
1689 oSerProProThr---ThrThrMetThrThrProSerProThrThrThrThrProSerPr 1708
QY TACCGGTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCGCACCCGGAATA 1477
Db : : : : : |||
1708 oile-----ThrThrThrThrThrThrProSerSerThrThrThrThrProSerProPr 1725
QY ACCTACCGGACGAAACCGGACCAAACTTTGGAAACGCGATCGCGCAAAAATCTCTATCC 1537
Db : : : : : |||
1725 oThrThrMetThrThrProSerProThrThrThrThrProSerProThrThrThrMetTh 1745
QY CTGCTGCCCA 1547
Db : : : : : |||
1745 rThrLeuPro 1748
RESULT 16
AAM24516
ID AAM24516 standard; protein; 5179 AA.
XX AAM24516;
AC
XX
XX 12-OCT-2001 (first entry)
XX
DE C899P predicted amino acid sequence.
XX Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
KW gene therapy; vaccine; colonic cancer.
XX
XX Homo sapiens.
XX
XX WO200149716-A2.
XX
XX 12-JUL-2001.
XX
XX 29-DEC-2000; 2000WO-US035596.
XX
XX 30-DEC-1999; 99US-00476296.
PR 10-JAN-2000; 2000US-00480321.
PR 15-FEB-2000; 2000US-00504629.
PR 06-MAR-2000; 2000US-00519444.
```

```
PR 19-MAY-2000; 2000US-00575251.
PR 29-JUN-2000; 2000US-00609448.
PR 28-AUG-2000; 2000US-00649811.
XX (CORI-) CORIXA CORP.
PA
XX
XX Xu J, Lodes MJ, Sclerist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX WPI; 2001-441847/47.
XX
PT Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer.
XX
PS Claim 2; Page 446-462; 472pp; English.
XX
CC The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity,
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be used
CC to treat disorders associated with decreased expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC TCAPs by expressing inactive proteins or to supplement the patients own
CC production of them. Additionally, (II) may be used to produce the TCAP
CC proteins, by inserting the nucleic acids into a host cell culturing the
CC cell to express the protein. (II) and its complementary sequences may
CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
CC and hybridisation assays to detect and quantitate the presence of similar
CC nucleic acids in samples, and therefore which patients may be in need of
CC restorative therapy. (I) may also be used as antigens in the production
CC of antibodies against TCAPs and in assays to identify modulators of TCAP
CC expression and activity. Anti-(I) antibodies and antagonists may also be
CC used to down regulate TCAP expression and activity. The anti-(I)
CC antibodies may also be used as diagnostic agents for detecting the
CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay
CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent
CC nucleotide and amino acid sequences given in the exemplification of the
CC present invention
XX
SQ Sequence 5179 AA;
Alignment Scores:
Pred. No.: 8.08e-16 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 4 Gaps: 22
US-10-665-990A-13 (1-1561) x AAM24516 (1-5179)
QY 119 ACGGAAAGCGCTCATTTCAAT-----ACTTCCAAACCTGTCTCTCTGGACAAC 166
Db : : : : : |||
1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGTGCAAAATCCGCGCACACCCCTCATAAACAGGGCTATCCGACATCTACTGTCTGAC 226
Db : : : : : |||
1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrThrTh 1287
QY 227 GACCCCGACGAGCCCTTCCGCCCGCCGCCCTTATCGAATCTGCCGACACAGCCTC 286
Db : : : : : |||
1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTCGAATACTACATTTGGCGCAACGACATTTCCGCGAGGCTGCTGTTCACACCTCA-- 344
Db : : : : : |||
1293 rValLeuSerThrThr-----ProlLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
```


CC development of cancer in a patient. (1) have immunostimulant and
 CC cytostatic activities and can be used in vaccines. AB232646 to AB233725
 CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
 CC sequences used in the exemplification of the present invention

XX Sequence 5179 AA;

Alignment Scores:

Pred. No.: 8.08e-16 Length: 5179
 Score: 336.50 Matches: 141
 Percent Similarity: 33.6% Conservative: 55
 Best Local Similarity: 24.1% Mismatches: 198
 Query Match: 11.8% Indels: 190
 DB: 6 Gaps: 22

US-10-665-990A-13 (1-1561) x ABP55365 (1-5179)

```

QY 119 ACGGAAAGCCGTCATTTCAT- - - - - ACTTCCAAACCTGTCTCTCTGGACAAC 166
D 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGTCAATCCGGCACACCCCTCATAAACACGGCTATCCGACATCTACCTGTCTGAC 226
D 1267 rPheThrThrIleThrLeuProThrProThrProThrSerPheThrThrThrThrTh 1287
QY 227 GACCCCAAGAGCCCTTGGCCGCCGCCCTTATCGAATCTGCCGAACACAGCCCTC 286
D 1287 rThrProThr - - - - - SerSerTh 1293
QY 287 GATTGTCAATACTACATTTGGCGCAACGACATTTCCGGCAGGCTGTCTCAACTCA- - 344
D 1293 rValLeuSerThrThr - - - - - ProlLysLeuCysCysLeuTrpSerAs 1307
QY 344 - - - - - 344
D 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
QY 345 - - - - - TGTACCTTGGCGAGAACGGCGGTGGCGGTACGGC- - - - - 380
D 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 - - - - - 380
D 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 - - - - - 380
D 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
QY 381 -TGCTGTGG- - - - - ACGACAAACAAC 400
D 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
QY 401 ACGCGCGGTGGAGCATCTCTGTCTCGCCCTCGACAGCATCCCAATATCGAAGTGGC 460
D 1407 rProProProThrThrThrThrThrThrLeuProProThrThrThrPro- - - - - 1422
QY 461 CTGTTCAACCCCTCGTCTACGCAAAATGGCGGCACCTCGGCTACCTGACCGACTTCCC 520
D 1423 - - - - - SerProProThrThrThrThrThrProProProThrThrThrProSerProPr 1441
QY 521 CGCCTCAACCGCGCATGACACAAATCTTTTACCGCGGACAAACGGCGGCACCATCTC 580
D 1441 oIleThrThr - - - - - ThrThrThrProLeuProThrThrThrProSerPro- - - - - 1456
QY 581 GCGCAGCGCAATATCGCGACGAATCTTCAAAGTCGGTGAGGACACCGTTTTTCGCCGAC 640
D 1457 - - - - - ProlLysSerThrThrThrThrProProProThrThrThrProSerProProTh 1474
QY 641 CTGGACATCTTCGCCACCGCGAGCGTGTGGCGAGTATCGCAGCACTTCGAC- - - - - 695
D 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494

```

```

QY 696 - - - - - GCTACTGGGCAAGCCATTTCGCCCAACAGCCGACCGCATCATCCGACGGC 748
D 1494 oProProThrThrThrProSerProMetThrThrProIleThrProProAlaSerTh 1514
QY 749 AACATCGGAGGCTCTTAAGCACTCGGATACAAACGAAACATCCAGACACCGCGCTC 808
D 1514 rThrThrLeu- - - - - ProProThrThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGCTACCGCGAAACCGTGAAC- - - - - AGTCGCCCTCTTACCAAAAAATACAGACG 862
D 1529 rThrThrProProThrThrThrProSerProProThrThr - - - - - 1544
QY 863 GGACGCTCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCAGACCCCTGCAAAAGGA 922
D 1545 - - - - - ThrProIleThrProPr 1550
QY 923 CTCGACCGCGACCGCGCAACCGCGATTGCGGGAGGCTGCAAGACGCGCTCAAAACAG 982
D 1550 oThrSerThrThrThr - - - - - LeuPro- - - - - ProThrThrThrProSe 1563
QY 983 CCCGAAAAAGCGTCTATCTGTTTCACCCCTATTTTCCTTACAAAATCCGGCAGAC 1042
D 1563 xProProProThrThrThrThrProProProThrThrThrProSerProProThrTh 1583
QY 1043 GCACGTGGCAAACTGGTGCAGGACGGCATAGACGTTACCGTCTGACCA- - - - - AC 1093
D 1583 r- - - - - ThrThrProSerProProThrIleThrTh 1593
QY 1094 TCGCTACAGGCGACGCGTTCGCGCTTCATTCGCGTACGTCAAT- - - - - 1142
D 1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 - - - - - ACGAAAAACGCTGTCTCAAGCGCGCATCAAACTC 1177
D 1613 oProThrThrThrProSerProProThrThrThrProIleThrProThrSerThrTh 1633
QY 1178 TACGAGCTGCAACCAACCATGCGTCCCGCCGCAACAAAGACAAAGGCTGACCGGACG 1237
D 1633 rThrLeuProProThrThrThrProSerProProProProThrThr - - - - - ThrThrTh 1650
QY 1238 TCCGTAAACGCGCTGCAATCCAAAACCTTCATTGTGGACGGCAACACGATCTTCATCGC 1297
D 1650 rProProProThrThrThrProSerProProThr- - - - - ThrThrThrProSerProPr 1669
QY 1298 TCATTCAACCTGACCCCGTTCGCGACGGCTCAATACCGAAATGGGCGTGTATCGAA 1357
D 1669 eThrThrThrThrProProProThrThrThrProSerSerProIleThrThrThrPr 1689
QY 1358 AGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGCATACCCACCCGAAATACGCC 1417
D 1689 oSerProProThr- - - - - ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
QY 1418 TACCGCGTTACCTCGACAAACACACACCGCTGCAATGGCAGCATCCCGCACCCGAAAA 1477
D 1708 oIle- - - - - ThrThrThrThrThrProSerSerThrThrThrProSerProPr 1725
QY 1478 ACCTACCGAAGCAACCGCAAGCAAACTTTGGAAACGCATCGCGCGCAAAATCTATCC 1537
D 1725 oThrThrMetThrThrProSerProThrThrThrProSerProThrThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
D 1745 rThrLeuPro 1748

```

RESULT 18

ABO07258
 ID ABO07258 standard; protein; 5179 AA.
 XX ABO07258;
 XX 13-AUG-2003 (first entry)
 XX Human p53 modifying protein, SEQ ID 218.

Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle; apoptotic disorder; cell proliferation disorder.

Homo sapiens.

WO200299122-A1.

12-DEC-2002.

03-JUN-2002; 2002WO-US017382.

05-JUN-2001; 2001US-0296076P.

10-OCT-2001; 2001US-0328605P.

15-FEB-2002; 2002US-0357253P.

(EXEL-) EXELIXIS INC.

Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

WPI; 2003-156859/15.

N-P5DB; ACD13431.

Identifying modulators of the p53 pathway for use in treating apoptotic or cell proliferation disorders, comprises screening for agents that modulate activity of a human ortholog of genes that modify the p53 pathway in *Drosophila*.

Example 2; Page 623-639; 678pp; English.

The invention relates to identifying (M1) a candidate p53 pathway modulating agent, by contacting an assay system comprising a purified HM polypeptide (human orthologue of genes that modify the p53 pathway in *Drosophila*) or nucleic acid with a test agent under conditions, where but for the presence of the test agent, the system provides a reference activity, and detecting a test agent-biased activity of the assay system. Also included are modulating (M2) a p53 pathway of a cell (comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a HM polypeptide comprising an HM amino acid sequence, where p53 function is restored), modulating (M3) a p53 pathway in a mammalian cell (comprising contacting the cell with an agent that specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4) a disease in a patient (comprising: (a) obtaining a biological sample from the patient; (b) contacting the sample with a probe for HM expression; (c) comparing the results with a control; and (d) determining whether the comparison indicates a likelihood disease). (M1) is useful for identifying modulators of the p53 pathway. A probe for HM expression is useful for diagnosing breast, colon, kidney, lung and ovarian cancer, in a patient, where the cancer has greater than 25 % expression level. Modulators identified by (M1) are useful in a variety of diagnostic and therapeutic applications, where disease or disorder prognosis is related to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell proliferation disorders (e.g. cancer). Another two new methods (M2 and M3) are useful for modulating the p53 pathway of a cell, thus restoring the p53 function of the cell, so that the cell undergoes normal proliferation or progression through the cell cycle. (M2) and (M3) are also useful for treating defects in the p53 pathway such as angiogenic, apoptotic or cell proliferation disorders. The present sequence represents a human p53 pathway modifying protein

Sequence 5179 AA;

Alignment Scores:

Pred. No.:	8.08e-16	Length:	5179
Score:	336.50	Matches:	141
Percent Similarity:	33.6%	Conservative:	55
Best Local Similarity:	24.1%	Mismatches:	198
Query Match:	11.8%	Indels:	190
DB:	6	Gaps:	22

US-10-665-990A-13 (1-1561) x ABO07258 (1-5179)

QY	119	ACGGAAGCGCTCATTTCAAT-----ACTTCCAAACCTGTCTCTCGACAAAC	166
Db	1248	ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh	1267
QY	167	ATCTGCAAAATCCGGCACACCCCTCATAAACAACGGGTATCCGACATCTACTGCTCGAC	226
Db	1267	rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrTh	1287
QY	227	GACCCACAGAGCCCTTGGCCGCGCCCTTATCGAATCTGCCGAACACAGCCTC	286
Db	1287	rThrProThr-----SerSerTh	1293
QY	287	GATTTGCAATATACATTGGCGCAACGACATTTCCGGCAGGCTGCTGTCAACCTCA--	344
Db	1293	rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs	1307
QY	344	-----	344
Db	1307	pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs	1327
QY	345	-----TGTAACCTTGGCGCAGAACGGCGTGGCGGTACGCC-----	380
Db	1327	pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe	1347
QY	380	-----	380
Db	1347	rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs	1367
QY	380	-----	380
Db	1367	nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs	1387
QY	381	-TGCTGTGG-----ACGACAAACAAAC	400
Db	1387	nCysCysTrpProMetAspLysCysIleThrThrProSerProThrThrThrProSe	1407
QY	401	ACGCGCGGTGGACGATCTCTCTCGCTCGCTCGACGACCATCCCAATATCGAAGTGGC	460
Db	1407	rProProThrThrThrThrThrLeuProThrThrPro-----	1422
QY	461	CTGTTCACCCCTTCTGTCTACGCAAAATGGCGGCACCTCGGCTACCTGACGACTCC	520
Db	1423	-----SerProProThrThrThrThrProProProThrThrThrProSerProPr	1441
QY	521	CGCCTCAACCGCGCATGCAACAATACTTACCGCGCAGCAACCGCGCCACCATCTC	580
Db	1441	oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro-----	1456
QY	581	GGCGGACGCAATATCGCGCAGCAATACTTCAAAGTCGGTGAGGACACCGTTTCCGCG	640
Db	1457	-----ProIleSerThrThrThrThrProProProThrThrThrProSerProThr	1474
QY	641	CTGACATCTCGCACCGGCGTGTGGCGGAAGTATCGCAGACTTCGAC-----	695
Db	1474	rThrThrProSerProProThrThrThrProSerProProThrThrThrThrPr	1494
QY	696	-----GCTACTGGGCAAGCCATTCCGCCCAACGACCGCGCATCATCCGCGCGGC	748
Db	1494	oProProThrThrThrProSerProMetThrThrThrProIleThrProProAlaSerTh	1514
QY	749	AACATCGCAAGGCTCTTCAAGCACTCGGATACAACGACAAACATCCAGACACGCGTC	808
Db	1514	rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh	1529
QY	809	CTCGCTACCGCAAAACCGTCGAAC-----AGTCGCCCTCTTACCAAAAAAATACAGC	862
Db	1529	rThrThrThrProProProThrThrThrProSerProProThrThr-----	1544
QY	863	GGACGATCATGCGGAGAGCGTCCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGA	922
Db	1545	-----ThrProIleThrProPr	1550

```
QY 923 CTCGACCGGACCGCGGCAAAACCGCGGATTTGCGGGAGGCTGCAAGACGCGCTCAAAACAG 982
Db : : : : :
1550 oThrSerThrThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 983 CCGGAAAAAGCGTCTATCTGGTTTCACCTATTTCGTCCTTACAAATCCGGCAGAC 1042
Db : : : : :
1563 rProProProThrThrThrThrProProThrThrProProProThrTh 1583
QY 1043 GCAGCTGGCAAAACTGGTGACGAGCGCATAGACGTTACCGTCTCGACCA-----AC 1093
Db : : : : :
1583 r-----ThrThrProSerProProThrThrThrThrThrThr 1593
QY 1094 TCGCTACAGCGACGACGCTTCCCGCGTCCATTCGGCTACGTCAAAT----- 1142
Db : : : : :
1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 -----ACGAAAAACCGCTCTCAAGCCGGCATCAAACTC 1177
Db : : : : :
1613 oProThrThrThrProSerProProThrThrThrThrProThrThrThrThrThrTh 1633
QY 1178 TAGGAGCTGCAACCCCAACCATCGCTCCCGCGCACAAAAAGAGCAAGGCGCTGACCGCAGC 1237
Db : : : : :
1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
QY 1238 TCGGTAAACGAGCTGATGCCAAAACCTTCATTGTGGAGCGCAAAAGCATCTTCATCGGC 1297
Db : : : : :
1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
QY 1298 TCATCAACCTGACCCCGTTCCGACGCTCAATACGAAATGGCGTCTCATCGAA 1357
Db : : : : :
1669 eThrThrThrThrThrProProProProThrThrThrThrProSerProThrThrThrPr 1689
QY 1358 AGCCCAAAATCGCAAGACAGATGGAGCGCACCTCGCGATACCACACCCGAATACGCC 1417
Db : : : : :
1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
QY 1418 TACCGGTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCGCCACCGCAAAA 1477
Db : : : : :
1708 oile-----ThrThrThrThrThrProSerSerThrThrThrThrProSerProPr 1725
QY 1478 ACCTACCGCAAGACCCGAAAGCCAACTTTGAAACGCATCGCCGCAAAAATCTCTATCC 1537
Db : : : : :
1725 oThrThrMetThrThrProSerProThrThrThrThrProSerProProThrThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
Db : : : : :
1745 rThrLeuPro 1748
RESULT 19
ADD48091
ID ADD48091 standard; protein; 5179 AA.
XX AC
XX ADD48091;
XX AC
XX 02-DEC-2004 (revised)
XX 29-JAN-2004 (first entry)
XX Human Protein NP_002448, SEQ ID NO 13789.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS
XX Homo sapiens.
XX Unidentified.
XX WO2003016475-A2.
XX PN
XX 27-FEB-2003.
XX PD
XX 14-AUG-2002; 2002WO-US025765.
XX PF
XX 14-AUG-2001; 2001US-0312147P.
XX PR
```

```
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX XX
PA (GEO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX XX
DR WPI; 2003-268312/26.
DR GENBANK; NP_002448.
XX XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX XX
PS Example 1; Page; 1017pp; English.
XX XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX XX
SQ Sequence 5179 AA;
Alignment Scores:
Pred. No.: 8,08e-16 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 7 Gaps: 22
US-10-665-990A-13 (1-1561) x ADD48091 (1-5179)
QY 119 ACGGAAAGCGCTCATTTCAAT-----ACTTCAAACCTGTCTCTCTGGACAAC 166
Db : : : : :
1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGCGAAATCCGCGACACCCCTCATACACGGGCTATCCGACATCTACCTGCTCGAC 226
Db : : : : :
1267 rPheThrThrThrLeuThrLeuProThrThrProThrSerThrThrThrThrThrTh 1287
QY 227 GACCCCGCAAGACCGCTTTCGCGCGCGCCCTTATCGAATCTCGCGAACACAGCCTC 286
Db : : : : :
1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTGCAATACTACATTTGGCGGCAACGACATTTCCGCGAGCTGCTGTTCAACCTCA-- 344
Db : : : : :
1293 rValLeuSerThrThr-----ProlLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
```

Db	1307	pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs	1327
Qy	345	-----TGTTACTTGGCGCAGAAACGGCGGTGGCGGTACGCC-----	380
Db	1327	pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe	1347
Qy	380	-----	380
Db	1347	rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs	1367
Qy	380	-----	380
Db	1367	nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs	1387
Qy	381	-TGCTGTTGG-----	400
Db	1387	nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrThrProSe	1407
Qy	401	ACGGCGGGTGGAGCATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGC	460
Db	1407	rProProProThrThrThrThrLeuProProThrThrPro-----	1422
Qy	461	CTGTTTCAACCCCTTCGTCCTACGCAATGGCGGCACCTCGGCTACCTGACCGACTTCCC	520
Db	1423	-----SerProProThrThrThrThrThrThrProProProThrThrThrProSerProPr	1441
Qy	521	CGCTCAACCGCGCATCGACAACAAATCTTTACC CGCGACACCGCGCCACATCATC	580
Db	1441	oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro-----	1456
Qy	581	GGCGGACGCAATATCGCGAGCAATACTTCAAAAGTCGGTAGGACACCGTTTTGCGCGAC	640
Db	1457	-----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh	1474
Qy	641	CTGGACATCTCGCCACCGGCGGTGGTGGCGAAGTATCGACGACTTCGAC-----	695
Db	1474	rThrThrProSerProProThrThrThrThrProSerProProThrThrThrThrThrPr	1494
Qy	696	-----GCTACTGGGCAAGCCATTCGCGCCCAACAAACCGCGCATCATCCGCGCGGC	748
Db	1494	oProProThrThrThrProSerProProMetThrThrProIleThrProProIleSerTh	1514
Qy	749	AACATCGGCAAGGTCTTCAAGCACTCGGATACAAGACGAAACATCCGACACGCGCTC	808
Db	1514	rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh	1529
Qy	809	CTGGCTACCGGNAACCGTCGAAC-----AGTCGCCCTCTACCAAAAATACAGACG	862
Db	1529	rThrThrThrProProProThrThrThrThrProSerProProThrThr-----	1544
Qy	863	GGACGCATCAGCTGGCAGAGCGTCCAAACCCGCTGTATGACGCAACCCCTGCAAAAGGA	922
Db	1545	-----ThrProIleThrProPro-----	1550
Qy	923	CTGACCGCGACCGCGCAAAACCGCGATTGCGGGAGGCTGCAAGACGCGCTCAACACAG	982
Db	1550	oThrSerThrThrThr-----LeuPro-----ProThrThrThrProSe	1563
Qy	983	CCCGAAAAAGCTCTATCTGGTTTACCCTATTTCGTCCCTACAAATCCGCGCAGAC	1042
Db	1563	rProProProThrThrThrThrThrProProProThrThrThrProSerProProThrTh	1583
Qy	1043	GCACTGGCAAACTGGTGCAGGCGGATAGAGTTACCGTCTCTGACCA-----AC	1093
Db	1583	r-----ThrThrProProProThrThrThrThrThrThrThrThrThrThrTh	1593
Qy	1094	TCGCTACAGCGACCGAGTTGCGCGCGTCCATTCCGGCTACGCTCAAAAT-----	1142
Db	1593	rThrThrProProProThrThrThrProSerProProThrThrThrThrThrThrProPr	1613
Qy	1143	-----ACGGAAAAACCGTGTGTCTCAAGCGCGCATCAAACTCAAACTC	1177

Db	1613	oProThrThrThrThrProSerProProThrThrThrThrProIleThrThrProThrSerThrTh	1637
Qy	1178	TACGAGCTGCAACCCACCAACCATGCGTCCCGCCACCAAAACACAAAGCGCTGACCGGCAGC	1237
Db	1633	rThrLeuProProThrThrThrThrProSerProProProThrThr-----ThrThrTh	1650
Qy	1328	TCCGTAAACGAGCTGCAATGCGAAACCTTCATTTGTGGACGGCAACGCATCTTCATCGGC	1297
Db	1650	rProProProThrThrThrThrProSerProProThr--ThrThrThrProSerProProIle	1669
Qy	1298	TCATTCAACTTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCGTCGTCATCGAA	1357
Db	1669	eThrThrThrThrThrProProThrThrThrThrThrProSerProIleThrThrThrPr	1689
Qy	1358	AGCCCCAAAATCGCAGAACAGATGGAGCGACCCCTCGCCGATACCAACCCCGAATAGCC	1417
Db	1689	oSerProProThr---ThrThrMetThrThrThrProSerProThrThrThrProSerPr	1708
Qy	1418	TACCGCTTACCTCGACAAACACACCGCTCGAATGGCAGCATCCGCGCACCCGAAAA	1477
Db	1708	oIle-----ThrThrThrThrThrProSerSerThrThrThrThrProSerProPr	1725
Qy	1478	ACCTACCCGAACGACCCGGAAGCCAAACTTTGAAACGCATCCCGCAAAAATCCCTATCC	1537
Db	1725	oThrThrMetThrThrProSerProThrThrThrThrProSerProProThrThrMetTh	1745
Qy	1538	CTGCTGCCCA 1547	
Db	1745	rThrLeuPro 1748	
RESULT 20			
ADD44998			
ID	ADD44998 standard; protein; 5179 AA.		
AC	ADD44998;		
DT	29-JAN-2004 (first entry)		
DE	Human Protein Q02817, SEQ ID NO 10430.		
KW	Human; pain; neuronal tissue; gene therapy;		
KW	spinal segmental nerve injury; chronic constriction injury; CCI;		
KW	spared nerve injury; SNI; Chung.		
OS	Homo sapiens.		
PN	W02003016475-A2.		
PD	27-FEB-2003.		
PF	14-AUG-2002; 2002WO-US025765.		
PR	14-AUG-2001; 2001US-0312147P.		
PR	01-NOV-2001; 2001US-0346382P.		
PR	26-NOV-2001; 2001US-0333347P.		
XX	(GEO) GEN HOSPITAL CORP.		
PA	(FARB) BAYER AG.		
XX	Woolf C, D'urso D, Befort K, Costigan M;		
XX	WPI; 2003-268312/26.		
DR	GENBANK; Q02817.		
XX	New composition comprising two or more isolated polypeptides, useful for		
PT	preparing a medicament for treating pain in an animal.		
XX	Claim 1; Page; 1017pp; English.		
XX	The invention discloses a composition comprising two or more isolated rat		
CC	or human polynucleotides or a polynucleotide which represents a fragment,		
CC	derivative or allelic variation of the nucleic acid sequence. Also		
CC	claimed are a vector comprising the novel polynucleotide, a host cell		

CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5179 AA;

Alignment Scores:

Pred. No.: 8,08e-16 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 7 Gaps: 22

US-10-665-990A-13 (1-1561) x ADD44998 (1-5179)

```
QY 119 ACAGAAAGCGCTATTTCAT-----ACTTCCAAACCTGCTCTCTGGACAC 166
    ||| :|||:|||||
    |||:|||||:|||||
Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
    ||| :|||:|||||
QY 167 ATCTGCAAAATCGGCACACCCCTCATACACAGGGCTATCCGACATCTACCTGCTCGAC 226
    ||| :|||:|||||
    |||:|||||:|||||
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrTh 1287
    ||| :|||:|||||
QY 227 GACCCCGACAGACCTTCCCGCGCGCGCCCTTATCGAATCTGCCGAAACACAGCCTC 286
    ||| :|||:|||||
    |||:|||||:|||||
Db 1287 rThrProThr-----SerSerTh 1293
    ||| :|||:|||||
QY 287 GATTTCGCAATACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCA-- 344
    ||| :|||:|||||
    |||:|||||:|||||
Db 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
    ||| :|||:|||||
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACTCTGCGCAGAACCGCGCGTGGCGGTACGCC----- 380
    ||| :|||:|||||
    |||:|||||:|||||
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
    ||| :|||:|||||
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
QY 381 -TGCTGTTGG-----ACGACAAAC 400
    ||| :|||:|||||
    |||:|||||:|||||
Db 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrThrProSe 1407
    ||| :|||:|||||
QY 401 ACGCGCGGGTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCG 460
```

```

    :||| :||| :||| :||| :|||
1407 rProProProThrThrThrThrThrLeuProProThrThrThrPro----- 1422
QY 461 CTGTTCAACCCCTCGTCTCTACGCAATAGGGCGGCGACTCGGTACCTGACCGACTTCCCC 520
    ||| :|||:|||||
    |||:|||||:|||||
Db 1423 ----SerProProThrThrThrThrThrProProProThrThrThrProSerProPr 1441
    ||| :|||:|||||
    |||:|||||:|||||
QY 521 CGCTCAACCGCGCATGCACAAATTCCTTTACCGCGGACAAACCGCGGCACCATCTC 580
    :||| :|||:|||||
    |||:|||||:|||||
Db 1441 oIleThrThr-----ThrThrProLeuProThrThrThrProSerPro----- 1456
    ||| :|||:|||||
    |||:|||||:|||||
QY 581 GCGGACCAATATCGCGAGCAATCTTCAAGTCGGTGAGGACACCGTTTTTCGCCGAC 640
    ||| :|||:|||||
    |||:|||||:|||||
Db 1457 -----ProIleSerThrThrThrThrProProThrThrThrProSerProProTh 1474
    ||| :|||:|||||
    |||:|||||:|||||
QY 641 CTGGACATCTCGCCACCGCGAGCGTCTGTCGGCGAAGTATCGCAGACTTCGACC----- 695
    ||| :|||:|||||
    |||:|||||:|||||
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrPr 1494
    ||| :|||:|||||
    |||:|||||:|||||
QY 696 -----GCTACTGGGCAAGCATTCCGCCACAAACGCGCAGCGCATCATCCGACGGC 748
    ||| :|||:|||||
    |||:|||||:|||||
Db 1494 oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh 1514
    ||| :|||:|||||
    |||:|||||:|||||
QY 749 AACATCGGCAAGGCTCTTCAAGCACTCGGATACACAGCAAGAAATCCAGACACGCGCTC 808
    ||| :|||:|||||
    |||:|||||:|||||
Db 1514 rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh 1529
    ||| :|||:|||||
    |||:|||||:|||||
QY 809 CTGCGCTACCGGAAACCGTGCAGC-----AGTCCGCCCTCTACCAAAAAATACAGACG 862
    :||| :|||:|||||
    |||:|||||:|||||
Db 1529 rThrThrThrProProProThrThrThrProSerProProThrThr----- 1544
    ||| :|||:|||||
    |||:|||||:|||||
QY 863 GGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGGACACCGCTCCCAAGGA 922
    ||| :|||:|||||
    |||:|||||:|||||
Db 1545 -----ThrProIleThrPro 1550
    ||| :|||:|||||
    |||:|||||:|||||
QY 923 CTCGACCGGACCGCGCAACCGCGGATTTGCGGGAGCGCTCGAAGCGCGTCAAAACAG 982
    ||| :|||:|||||
    |||:|||||:|||||
Db 1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
    ||| :|||:|||||
    |||:|||||:|||||
QY 983 CCGCAAAAAAGCGTCTATCTGGTTTCAACCTATTTTCGTCCTTACAAAATCCGGCAGAC 1042
    ||| :|||:|||||
    |||:|||||:|||||
Db 1563 rProProProThrThrThrThrProProProThrThrProSerProProThrTh 1583
    ||| :|||:|||||
    |||:|||||:|||||
QY 1043 GCATCGGCAAAACTGGTGACGACCGCATAGACGTTACGTCCTGACCA-----AC 1093
    ||| :|||:|||||
    |||:|||||:|||||
Db 1583 r-----ThrThrProSerProProThrIleThrTh 1593
    ||| :|||:|||||
    |||:|||||:|||||
QY 1094 TCGCTACAGGCGACGAGTTCGCCCGCTCCATTCGGGTACGCTCAAAAT----- 1142
    ||| :|||:|||||
    |||:|||||:|||||
Db 1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrProPr 1613
    ||| :|||:|||||
    |||:|||||:|||||
QY 1143 -----ACGAAACCGCTGCTCAAAGCGCGCATCAAACTC 1177
    ||| :|||:|||||
    |||:|||||:|||||
Db 1613 oProThrThrThrProProProThrThrProIleThrProProThrSerThrTh 1633
    ||| :|||:|||||
    |||:|||||:|||||
QY 1178 TAGGAGCTGCACCAACCATGCTCCCGCCGCCCAAAAAGACAAAGGCTGACCGCGACG 1237
    :||| :|||:|||||
    |||:|||||:|||||
Db 1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
    ||| :|||:|||||
    |||:|||||:|||||
QY 1238 TCGGTAAACGAGCTCATGCCAAAACCTTTCATTGTGGACGGCAACCGCATCTTCATCGCG 1297
    ||| :|||:|||||
    |||:|||||:|||||
Db 1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
    ||| :|||:|||||
    |||:|||||:|||||
QY 1298 TCATTCAACCTCGACCCCGTTCCGACGGCTCAATACGGAATGGCGGCTCGTCATCGAA 1357
    :||| :|||:|||||
    |||:|||||:|||||
Db 1669 eThrThrThrThrThrProProProThrThrThrProSerProIleThrThrThrPr 1689
    ||| :|||:|||||
    |||:|||||:|||||
QY 1358 AGCCCAAAATCGCAGACAGATGAGCGCACCCCTCGCGATACACACACCCGAAATACGCC 1417
    :||| :|||:|||||
    |||:|||||:|||||
Db 1689 oSerProProThr-----ThrMetThrThrProSerProThrThrThrProSerSerPr 1708
    ||| :|||:|||||
    |||:|||||:|||||
QY 1418 TACCGGTATACCTCGACAAAACACACACCGCTTCAATGGCATGCGACGATCCCGCCCAAAA 1477
    :||| :|||:|||||
    |||:|||||:|||||
```



```
Db 1407 rProProThrThrThrThrThrLeuProProThrThrPro----- 1422
Qy 461 CTGTTCAACCCCTTCGTCTACGCAATGGCGGCACCTCGGTACCTACCGACTTCC 520
Db 1423 ---SerProProThrThrThrThrThrProProThrThrProSerProPr 1441
Qy 521 GCGCTCAACCGCGCGATGACACAAATCTTTACCGCGACACACCGCGCCACCATCTC 580
Db 1441 oileThrThr-----ThrThrProLeuProThrThrProSerPro----- 1456
Qy 581 GCGGACGCAATATCGGCGACGAATACTTCAAGTCGGTGAGGACACCGTTTCCCGCAC 640
Db 1457 -----ProLeuSerThrThrThrThrProProThrThrThrProSerProTh 1474
Qy 641 CTGGACATCTCGGCACCGCGAGCTGTCGCGAAGATGCGACGACTTCGACC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProThrThrThrThrPr 1494
Qy 696 -----GCTACTGGGCAAGCCATTCCGCCACACACGCGCATCATCCGACGCGC 748
Db 1494 oProProThrThrThrProSerProMetThrThrProLeuThrProAlaSerTh 1514
Qy 749 AACATCGGCAAGGTCCTCAAGCATCGGATACAGACGAAACATCCAGACGCGCTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrProSerProThrThrTh 1529
Qy 809 CTGCGCTACCGGAACCGTCGAAC-----AGTCGCCCTCTACCAAAATAACAGACG 862
Db 1529 rThrThrThrProProThrThrThrProSerProThrThr----- 1544
Qy 863 GGACGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGA 922
Db 1545 -----ThrProLeuThrProPr 1550
Qy 923 CTCGACCGCGACCGCGCAACCGCGATTGCGGGAGGTCGAGACGCGCTCAACAG 982
Db 1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
Qy 983 CCCGAAAGGCTATCTGTTTACCCCTATTTCGTCCTCAAAATCCGGCACAGC 1042
Db 1563 rProProThrThrThrThrThrProProThrThrThrProSerProThrTh 1583
Qy 1043 GCACTGGCAAACTGGTGACGAGCGCATAGACGTTACCGTCTCGACCA-----AC 1093
Db 1583 r-----ThrThrProSerProProThrThrThrThrTh 1593
Qy 1094 TCGCTACAGGCGACGAGCTTGGCGCGCTCATTCCGGCTACGTCAAT----- 1142
Db 1593 rThrThrProProThrThrThrProSerProThrThrThrThrThrThrProPr 1613
Qy 1143 -----ACGAAACCGCTGCTCAAGCGCGCATCAAACTC 1177
Db 1613 oProThrThrThrProSerProProThrThrThrProLeuThrProThrThrTh 1633
Qy 1178 TACGAGCTGCAACCAACCTCGCTCCCGCCCAAAAGACAAAGGCGCTACCGCGACG 1237
Db 1633 rThrLeuProProThrThrThrProSerProProThrThr-----ThrThrTh 1650
Qy 1238 TCGGTAAACCGCTGCATGCGAAACCTTCAATGTGGACGCGCAACGATCTTCATCGCG 1297
Db 1650 rProProThrThrThrThrProSerProThr-----ThrThrThrProSerProPr 1669
Qy 1298 TCATTCACTCGACCCCGTTCGCGAGGCTCAATACCGAAATGGGCGTGCATCGAA 1357
Db 1669 eThrThrThrThrThrProProThrThrThrThrProSerSerProLeuThrThrTh 1689
Qy 1358 AGCCCCAAATCGCAGAACAGATGAGCGCACCTCGCCGATACCAACACCGCAATAGCC 1417
Db 1689 oSerProThr---ThrThrMetThrThrProSerProThrThrThrThrProSerPr 1708
Qy 1418 TACCGGTTTACCTCGCAACACACACGCGCTGCAATGGCAGCATCCGCGACCCGAAA 1477
```

```
Db 1708 oile-----ThrThrThrThrThrProSerSerThrThrThrProSerProPr 1725
Qy 1478 ACCTACCGACGACCGAACCGCAACTTTGGAAACGATCGCGCGCAAAATCCTATCC 1537
Db 1725 oThrThrMetThrThrProSerProThrThrThrProSerProThrThrMetTh 1745
Qy 1538 CTGCTGCGCCA 1547
Db 1745 rThrLeuPro 1748

RESULT 23
AEA27639
ID AEA27639 standard; protein; 5179 AA.
XX
AC AEA27639;
XX
DT 28-JUL-2005 (first entry)
XX
DE Human intestinal mucin.
XX
KW Colon tumor; breast tumor; pseudomyxoma peritonei; prostate cancer;
KW andrology; genitourinary disease; neoplasm; endocrine disease;
KW gynecology and obstetrics; gastrointestinal disease; cytostatic; mucin;
KW RNA interference; RNAi; gene silencing; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2005047321-A2.
XX
PD 26-MAY-2005.
XX
PF 15-NOV-2004; 2004WO-US038116.
XX
PR 14-NOV-2003; 2003US-0520479P.
PR 23-JAN-2004; 2004US-0538740P.
PR 04-JUN-2004; 2004US-0577002P.
XX
PA (UYCR-) UNIV CREIGHTON.
XX
PI Loggie BW, Nawaz Z;
XX
DR WPI; 2005-367000/37.
XX
DR N-PSDB; AEA27625.
XX
PT New RNA polynucleotide, encoding a Ubch7, E6-AP, mucin-1, or mucin-2
PT polypeptide, useful for treating cancer, e.g. breast, prostate, or colon
PT cancer.
XX
PS Claim 4; SEQ ID NO 47; 94pp; English.
XX
CC The present invention provides polynucleotides that inhibit expression of
CC polypeptides such as the ubiquitin conjugating enzyme Ubch7, the
CC ubiquitin protein ligase E6-AP, mucin-1 or mucin-2. The polynucleotides
CC may be RNA polynucleotides that include a sense strand and an antisense
CC strand. The sense strand of the RNA polynucleotide comprises a sequence
CC of 16-30 nucleotides that is substantially identical to consecutive
CC nucleotides of an mRNA encoding the target polypeptide. The
CC polynucleotide may be administered to a cell in vivo or in vitro. The
CC administration can be to a subject, preferably a human, and can occur
CC during surgical resection of a tumor or resection of mucinous material.
CC Claimed methods for treating cancer include administering a
CC polynucleotide, preferably an RNA polynucleotide, of the invention to a
CC subject that has, or is at risk for, an endocrine-related cancer, such as
CC breast cancer or prostate cancer, or a cancer comprising mucin
CC overproduction, such as colon cancer, breast cancer or pseudomyxoma
CC peritonei, so that a symptom associated with the cancer is decreased. The
CC present sequence is that of human intestinal mucin. mRNA for MUC2
CC represents a target for gene silencing using RNA polynucleotides of the
CC invention AEA27626, AEA27627 and AEA27630.
XX
SQ Sequence 5179 AA;
XX
Alignment Scores:
```

Pred. No.: 8,08e-16 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: Gaps: 22

US-10-665-990A-13 (1-1561) x AEA27639 (1-5179)

QY 119 ACGGAAAGCCGTCATTCAAT-----ACTTCAAACCTGTCTCTCGGACAC 166
Db :|||:|||||
1248 ThrValGluLysHisPheAsnLeuCysserIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGCAAAATCGGCACACCCCTCATACACAGGGCTATCGACATCTACCTGCTCGAC 226
Db :|||:|||||
1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrTh 1287
QY 227 GACCCCAAGCCCTTCCCGCCGCGCCCTTATCGAATCTGCGCAACACAGCCTC 286
Db :|||:|||||
1287 rThrProThr-----SerThr 1293
QY 287 GATTGCAATACTACATTTGGCGCAACGACATTTCCGCGAGCTGTGTCAACCTCA-- 344
Db :|||:|||||
1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTipSerAs 1307
QY 344 ----- 344
Db :|||:|||||
1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACTTCCGCGCAGAACGCGCGGTGCGGTACGCC----- 380
Db :|||:|||||
1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db :|||:|||||
1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db :|||:|||||
1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
QY 381 -TGCTCTTGG-----ACGACACAAAC 400
Db :|||:|||||
1387 nCysCysTyrProMetAspLysCysIleThrThrProSerProProThrThrThProSe 1407
QY 401 ACGCGCGGTGGACGATCTCTGCTCGCCTCGACAGCCATCCCAATATCGAAGTGCOC 460
Db :|||:|||||
1407 rProProProThrThrThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTCGTCTCAGCAATGCGCGCACTCGGCTACTGACCGACTTCCGC 520
Db :|||:|||||
1423 ----SerProProThrThrThrThrThrThrThrThrThrThrThrThrThrProPr 1441
QY 521 CGCCTCAACCGCGCATGCACAAATCTTTACCGCGCAGCAACCGCGCCACCATCTC 580
Db :|||:|||||
1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGGAGCAATATCGCGACCAATATCTTCAAAGTCGGTGAGGACACCCGTTTTTCGCGCAG 640
Db :|||:|||||
1457 -----ProIleSerThrThrThrThrThrThrThrThrThrThrThrThrProProTh 1474
QY 641 CTGGACATCTCGCCACCGCGCAGCTGCTCGGGAAGTATCGCAGACTTCGACC----- 695
Db :|||:|||||
1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
QY 696 -----GCTACTGGCAAGCATTCGCGCCACACAGCCAGCGCATCATCGCGAGCGGC 748
Db :|||:|||||
1494 oProProThrThrThrProSerProProMetThrThrProIleThrProProIleSerTh 1514
QY 749 AACATCGGAAGGCTTCAAGCACTCGGATACACGACGAACATCTCAGACAGCGGCTC 808
Db :|||:|||||
1514 rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh 1529

QY 809 CTGCGCTACCGGAAACCGTCGAAC-----AGTCGCCCTCTACCAAAAATACAGACG 862
Db :|||:|||||
1529 rThrThrProProProThrThrThrProSerProProThrThr----- 1544
QY 863 GGACGCATCGACTGGCAGAGCGTCAAAACCGCCCTGATCAGCGACACCCCTGCAAAAGGA 922
Db :|||:|||||
1545 -----ThrProIleThrProPr 1550
QY 923 CTCGACCGGACCGCCGCAAAACCGCCGATTTGCCGGAGGCTGCAAGAGCGGCTCAAAACAG 982
Db :|||:|||||
1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 993 CCGGAAAAAGCGTCTATCTGTTTCACCCCTATTTGCTCCCTACAAAATCGGGCACAGAC 1042
Db :|||:|||||
1563 rProProProThrThrThrThrThrProProProThrThrThrProSerProProThrTh 1583
QY 1043 GCACCTGGCAAAACTGGTCGAGCAGCGCATAGACGTTTACCGTCTCTGACCA-----AC 1093
Db :|||:|||||
1593 r-----ThrThrProSerProProThrIleThrTh 1593
QY 1094 TCGCTACAGGCGACCGAGCTTGGCCCGCTCCATTCGGCTAGCTCAAT----- 1142
Db :|||:|||||
1593 rThrThrProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 -----ACGAAAAACCGTCTCAAAAGCGCGCATCAAACTC 1177
Db :|||:|||||
1613 oProThrThrThrProSerProProThrThrThrProIleThrProThrSerThrTh 1633
QY 1178 TACGAGCTGCAACCCCAACCATGCTCCCGCCACCAAAAGCAAAAGCGCTGACCGGCAGC 1237
Db :|||:|||||
1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
QY 1238 TCGGTAAACAGCCTCATGCCAAAACCTTCATTTGTGGAGCGCAACGATCTTCATCGGC 1297
Db :|||:|||||
1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
QY 1298 TCATTCAACCTGACCCCGTTCGCGCAGCGCTCAATACCGAAATGGCGTCTCATCGAA 1357
Db :|||:|||||
1669 eThrThrThrThrProProProThrThrThrThrProSerProIleThrThrThPr 1689
QY 1358 AGCCCCAAAATCGCAGAACAGATGGAGCGCACCTTCGCGCATACACACCCGAAATACGCC 1417
Db :|||:|||||
1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
QY 1418 TACCGCGTTCCTCGACAAACAGACCGCTGCAATGGCAGATCCCGCCACCGCAAAA 1477
Db :|||:|||||
1708 oIle-----ThrThrThrThrProSerSerThrThrThrThrProSerProPr 1725
QY 1478 ACCTACCGAAGCAACCGCAACCAACTTTGGAAACGATCGCGCAAAAATCTCTATCC 1537
Db :|||:|||||
1725 oThrThrMetThrThrProSerProThrThrThrProSerProProThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
Db :|||:|||||
1745 rThrLeuPro 1748

RESULT 24
ID ABUS3160
XX ABUS3160 standard; protein; 395 AA.
XX AC ABUS3160;
XX XX
XX 14-APR-2003 (first entry)
XX DE Human testes-derived DKFzphes3_2a11 homologue #20.
XX DE Human; gene therapy; vaccine; disease treatment; detection.
XX OS Homo sapiens.
XX PN WO200112659-A2.
XX XX
XX 22-FEB-2001.


```
QY 1039 AGACGACTGGCAAACTGGTCAGGACGGCATAGACGTTACCGTCCTGACCAACTCGCT 1098
Db 467 -----AlaaspHisArgArgArgProArgGlyGlnArg 478
QY 1099 ACAGCGACCGACGTTGGCGCGTCCTCATTCGGCTACGT----- 1137
Db 479 ArgGluSerGluGlnLeuArgHisProGlyArgArgGlnProAlaIleGluArgAla 498
QY 1138 -----CAATACCGAAACCGCTGCTCAACGCGG----- 1167
Db 499 ArgProAlaGlnArgAlaArgValAlaValGlnTrpArgArgHisAspProAlaArgArg 518
QY 1168 -----CATCAACTCTACGAGCTGCAACCCCAACCA-----TGCGTCCCCGCG 1209
Db 519 ArgHisHisGlnProLeuArgGlyCysValGlnProArgGlnAlaGlyTyrArgProArg 538
QY 1210 CACAAAGACAA-----AGCCTGACCGGAGCTCCGT 1242
Db 539 ArgCysArgGlnProCysArgGlnProAlaGlnProPheArgGlyAspArgGluArgGln 558
QY 1243 AAC---CAGCTGTCATGCCAA-----1260
Db 559 GlyLeuGlnProAlaCysGlnProAspArgGluProSerArgArgAlaGlyLysGlnVal 578
QY 1261 -----AACCTTCAT-----1269
Db 579 GlyProValHisArgGluAspGlyAlaAspArgLeuHisArgArgGlnArgGlyArg 598
QY 1269 -----1269
Db 599 LeuGlnArgGlnAlaGlnArgHisLeuAspHisProAlaArgGlnAspArgGlyHis 618
QY 1270 -----TGTGACGGCAACCGCATCTTCATCGCTCATTTCAA 1305
Db 619 ArgGlnGlnArgHisGlyAlaThrAlaGlyArgArgLeuArgHisArgArgArgHis 638
QY 1306 CTGACCCCGCTTCGCGACGGCTCAATACCGAAATGGCGTCTGCATCGAAAGCCCA 1365
Db 639 ProGluGlnProPheGlnSerAspArgGlnArgArg---GlnProHisArgGlnProArg 657
QY 1366 -----AATCGCAGACAGATGGAGCGCACCTCGC---CGA 1398
Db 658 SerProArgGlnProGlyProGlyAsnArgAla---GlyAsnHisProArgAlaAla 676
QY 1399 TACCACACCGCAATACGCTACCGGTTAC-----1428
Db 677 TyrArgSerArgArgArgTyrArgArgHisArgProGluValAlaGlnLeuHisGlnPro 696
QY 1429 -----CTCGACAAACACAAACCGCTGCAATGGCACGA-----1461
Db 697 LeuLeuValProGluArgGlnPheArgProGlyAlaArgGlyArgAspProArgArgAla 716
QY 1462 TCCGCGCACCGCAAAACCTA-----CCCGAA-----1488
Db 717 GlnArgAspProGlnArgLeuValLeuArgValArgIleProGluGlnGlyAlaAspPro 736
QY 1489 -----CGAACCGGAAGC 1500
Db 737 AspGlnGlnTrpArgProValLeuArgSerGlyAspProGlyArgArgArgHisGly 756
QY 1501 CAAACTTTGGAAGCGATCGCGCAAAATCTTATCCCTGCTGCCATCGAAGG 1554
Db 757 GlnCysGlnHisAlaHisArgGlnArgHisProProArgLeuHisLeuArg 774
RESULT 26
ABU53159
ID ABU53159 standard; protein; 386 AA.
XX AC
XX AC
XX ABU53159;
XX 14-APR-2003 (first entry)
XX
```

Human testes-derived DKFZphtes3_2a11 homologue #19.

Human; gene therapy; vaccine; disease treatment; detection.

Homo sapiens.

WO200112659-A2.

22-FEB-2001.

18-AUG-2000; 2000MO-IB001496.

18-AUG-1999; 99US-0149499P.

28-SEP-1999; 99US-0156503P.

(GEHU-) GERMAN HUMAN GENOME PROJECT.

Wiemann S;

WPI; 2001-327840/34.

Nucleic acids having the sequences of clones isolated from libraries of different human tissues, useful in recombinant DNA methodologies.

Example III; Page 774; 1095pp; English.

This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence represents a homologue of a polypeptide described in the disclosure of the invention

Sequence 386 AA;

Alignment Scores:

Pred. No.: 3,03e-15 Length: 386

Score: 325.00 Matches: 140

Percent Similarity: 38.2% Conservative: 45

Best Local Similarity: 28.9% Mismatches: 169

Query Match: 11.4% Indels: 130

DB: 4 Gaps: 24

US-10-665-990A-13 (1-1561) x ABU53159 (1-386)

QY 148 ACCTGTCTCTCTGGACAAACATCTGTGCAAAATCCGSCACACCCCTCATACACGCGCTATC 207

Db 17 ThrLeuProProThrThrThrProSerProPro--ThrThrThrThrThrThrProProp 36

QY 208 CGACATCTACTGCTGCTG-----ACGACCCCGCAGAGCCCTTGCCG-----C 249

Db 36 roThrThrThrProSerProProIleThrThrThrThrProLeuProThrThrThrP 56

QY 250 CCGCGCGCGCTTATCGAATCTGCCGACACAGCCTCGATTGCAATACATACATTGGCG 309

Db 56 roSerProProIleSerThr-----ThrThrThrProProp 68

QY 310 CAACGACATTTCCGCGCAGGCTGCTGTTCAACCTCATGCTTCCCGCAGACACGCGGT 369

Db 68 roThrThrThrPro-----SerProProThrThrThrProSerProProThrT 84

QY 370 GCGCGTACGCTGCTGTTGGACGACACAAACACGCGCGGTTGGACGATCTCCTGCTCGC 429

Db 84 hrThrProSerProProThrThrThrThrThrThrThrThrThrThrThrThrThr 104

QY 430 CCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAAC-----CCTCGCTCT 480

Db 104 roProMetThrThrProIleThrProProAlaSerThrThrThrLeuProProThrThr 123

```
QY 481 ACGCAATGGCGGCACTCGGTACTGACCGACTTCCCGCCCTCAACCGCGCATGCA 540
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
124 -----ThrProSerProThrThrThrThrThrThrThrThrThrThrThrThr 138
QY 541 CAACAAATCCTTACCGCGCAACAC-----GCGCCACTACTCGCGGCGCAA 591
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 hrThrThrProSerProProThrThrThrProIleThrProProThrThrThrL 158
QY 592 TATCGCGCAGAAATCTCAAGTCGTTGAGGACACCGTTTTCGCGCAGTGCATCCT 651
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
158 euProProThrThrProSerPro-----ProProThrThrThrThrThr 173
QY 652 CGCCACCGCAGCTGTCGGCGAAGTATCGACGACTTCGACCGCTACTGGCGAAGCA 711
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
173 hrProProProThrThrThrProSerProProThrThrThrThrProSerProThrI 193
QY 712 TTCGCGCCCAACAGCGCATCATCGCAGCGGCAACATCGCGAAGGTCCTCAAGC 771
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193 leThrThrThrThrPro-----ProProThrThrThrThrThrThrThr 198
QY 772 ACTCGGATACAGCAAGAAATCCAGACACGCGTCTCGCTACCGCAAAACCGTGA 831
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
199 --ProProThrThrThrProSerProProThrThrThrThrThrThrThrThr 218
QY 832 AC-----AGTCGCCCTCTACCAAAATACAGAGCGGACGATCGACTGCGAGCGT 885
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 hrThrProSerProProThrThr-----ProProThrThrThrThrThr 225
QY 886 CCAACCGCGCTGATCAGCGACACCCCTGCAAAAGACTCGACCGGACCGCGCAAC 945
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
226 -----ThrProIleThrProProThrThrThrThrThrThrThrThrThr 236
QY 946 GCCGATTGCGGAGGCTCAAGACGCGCTCAACAGCCGCGCAAAAGCGTCTATCGT 1005
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 -----LeuPro-----ProThrThrThrProSerPro----- 245
QY 1006 TTCACCTATTTCGTCCTCCCTACAAATCCGCGACAGCAGCTGGCAAACTGGTGCAG 1065
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
246 -----ProProThrThrThrThrThrThrProProThr----- 256
QY 1066 CGGCATAGACGTACCGTCTGACCAACTTCGTCGACGCGGACCGGCGCGTCCA 1125
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 -----ThrThrProSerProProThr-----ThrThrThrProSerP 269
QY 1126 TTCGCGTACGTCAATACCGAAACCGCTGCTCAAGCGCGCATCAAACTACGAGCT 1185
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
269 roProIleThrThrThrThr-----ThrProProThrThrThrProS 284
QY 1186 GCAACCCA-----ACCATGCGTCCCGCCACAAAGACAAAGCGCTGACCGGAGCTC 1239
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284 exSerProIleThrThrThrThrProSerProProThrThrMetThrThrProSer 302
QY 1240 CGTAACACCGCTGCATGCCAAAACCTTCATTGTGGACGGCAACGCACTTTCATCGGCTC 1299
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
303 -----ProThrThrThrProSerSerProIleThrThrThrThrThrProSer 318
QY 1300 ATTCAACTCGACCCCGTTCGCGACGCGTCAATACCGAAATGGCGGTGCATCGAAG 1359
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
319 --SerThrThrThrProSerPro-----ProProThrThrMetThrThrProS 334
QY 1360 CCCCAGAAATCGGAGACAGATGGAGCGCACCTTCGCGGATACACACCGCAATAGCCTA 1419
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 exProThrThr-----ThrProSerProProThrThrThrThrMet 347
QY 1420 CGCGGTACCTTCGCAACACAAACCGCTGCAATGGCAGCATCCGCGCACCGCAAAAC 1479
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
347 hrThrLeuProProThrThrThrSerSerProLeuThrThrThrProLeuProSerI 367
QY 1480 CTACCGGAACGAAACCGAAGCAAACTTTGGAAACGCAATCGCGGCAAAATCCTATCCCT 1539
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
367 leThrProProThrPheSerPro---PheSerThrThrThrThrPro-----ThrThrProC 384
QY 1540 GCTGCCCA 1547
```

```
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
384 ysValPro 386
RESULT 27
ABU53144
ID ABU53144 standard; protein; 717 AA.
XX
AC ABU53144;
XX
DT 14-APR-2003 (first entry)
XX
DE Human testes-derived DKFZphtes3_2a11 homologue #4.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000MO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Example III; Page 765; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a homologue
CC of a polypeptide described in the disclosure of the invention
XX
SQ Sequence 717 AA;
Alignment Scores:
Pred. No.: 3,52e-15 Length: 717
Score: 325.00 Matches: 140
Percent Similarity: 38.2% Conservative: 45
Best Local Similarity: 28.9% Mismatches: 169
Query Match: 11.4% Indels: 130
DB: 4 Gaps: 24
US-10-665-990A-13 (1-1561) x ABU53144 (1-717)
QY 148 ACCTGCTCTCTCGGACAAACATCTGCAATCGGCACACCCCTCATACACAGCGGTATC 207
Db 15 ThrLeuProProThrThrThrProSerProPro--ThrThrThrThrThrThrPro 34
QY 208 CGACATCTACCTGCTCG-----ACGACCCCAACGAGCCCTTGCGG-----C 249
Db 34 roThrThrThrProSerProIleThrThrThrThrProLeuProThrThrThrP 54
QY 250 CGCGCGCGCCCTTATCGAATCGCGAACAACAGCCCTCGATTTGCAATCTACATTGGCG 309
Db 54 roSerProProIleSerThr-----ThrThrThrThrThrThrThrThrThr 66
QY 310 CAACGACATTTCCGGCAGGCTGCTGTCAACCTCATGTACCTTGGCGCAAGACGGCGCT 369
```


Db	219	hrThrProSerProThrThrThr	:::	-----	226
Qy	886	CCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACGCGACGCGCGAAACC	::: :::	-----	945
Db	227	-----ThrProIleThrProThrSerThrThr	-----	-----	237
Qy	946	GCGGATTCGCGGAGGCTGCAGACGCGGTCTCAACAGCGCGCAAAAAGCGTCTATCTGGT	::: :::	-----	1005
Db	238	-----LeuPro-----ProThrThrProSerPro	-----	-----	246
Qy	1006	TTCAACCTATTTCGTCCTTACAAANATCCGCACAGACGCACTGGCAAAACTGGTCAGGA	:::	-----	1065

Qy	1066	CGCATAGACGTTACCGTCTCTGACCAACTCGCTACAGGCGACCGAGCTTCCGCGCGTCCA	1125
Db	258	-----ThrThrProSerProProThr-----ThrThrThrProSerP	270
Qy	1126	TTCCGGCTACGTCAAATAACCGAAACCGCTGCTCAAAGCGCGCATCAAACTCTACGAGCT	1185
Db	270	roProlleThrThrThr-----ThrProProThrThrThrProS	285
Qy	1186	GCAACCCA-----ACATGCGCTCCCGGCCACAAAGACAAAGGCGTGAACCGGAGCTC	1239
Db	285	erSerProlleThrThrThrProSerProProThrThrThrMetThrThrProSer----	303
Qy	1240	CGTAACGAGCTGCATGCGCAAAACCTTCATTGTGGAGCGCAACGCGATCTTCATCGGCTC	1299
Db	304	-----ProThrThrThrProSerSerProleThrThrThrThrProSer-----	319
Qy	1300	ATTCAACTCGACCCCGTTCCGACGCGTCAATACGAAATGGGCGTCGTCAATCGAAAG	1359
Db	320	--SerThrThrThrProSerPro-----ProProThrThrMetThrThrProS	335
Qy	1360	CCCCAAATCGCAGAACAGATGGAGCGCACCCCTCGCGGATACACACCCCGAATACGCCTA	1419
Db	335	erProThrThr-----ThrProSerProProThrThrThrMetT	348
Qy	1420	CCGGTTTACCTCGACAAACACAAACCGCTGCAATGGCAGCATCCGCGCACCC	1472
Db	348	hrThrLeuProProThrThrSerSerProLeuThrThrThrProLeuPro	365
RESULT 31			
ABM71816			
ID	ABM71816 standard; protein; 513 AA.		
AC	ABM71816;		
DT	20-NOV-2003 (first entry)		
DE	Staphylococcus aureus protein #1056.		
KW	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;		
KW	enzymatic assay; antibiotic target.		
OS	Staphylococcus aureus.		
PN	WO200294868-A2.		
PD	28-NOV-2002.		
PF	27-MAR-2002; 2002WO-IB002637.		
PR	27-MAR-2001; 2001GB-00007661.		
PA	(CHIR-) CHIRON SPA.		
PI	Masignani V, Mora M, Scarselli M;		
DR	WPI; 2003-120786/11.		
DR	N-ESDB; ACF73376.		

PT New *Staphylococcus aureus* protein, useful as a vaccine for treating or
PT preventing *Staphylococcal* infection, specifically an infection caused by
PT *S. aureus*, e.g. sepsis.

The invention relates to novel genes and encoded proteins from *Staphylococcus aureus*. A composition comprising the *S. aureus* protein, a nucleic acid encoding the protein, or an antibody to the protein, is useful as a pharmaceutical, particularly as a vaccine for treating or preventing infection due to *Staphylococcus* bacteria, specifically an infection caused by *S. aureus*. The composition is particularly useful for treating or preventing sepsis in a patient. The composition can also be used for diagnostics. The protein is also used in an assay for enzymatic studies and as a target for antibiotics. This sequence represents one of the novel *S. aureus* proteins of the invention.

Sequence 513 AA;

Alignment Scores:		
Pred. No.:	1.14e-13	Length:
Score:	304.00	Matches:
Percent Similarity:	39.0%	Conservative:
Best Local Similarity:	24.3%	Mismatches:
Query Match:	10.7%	Indels:
DB:	6	Gaps:
		17
		513

US-10-665-990A-13 (1-1561) x ABM71816 (1-513)

	QY	134	TTCAAATACTTCCAAACCTGTCCTCTGTGGACAACATCTCGCAATCCGGCACACCCTCAT	191
Db		140	TyrAsnAlaIAPheLeuThrThraSPasnAepLeuLysIlle---TyrThrAspGly	158
QY		194	AACAACGGGTATCCGACATCTACTGCTGCAGCAGCCCCACGAAGACCTTGCCGCCCGC	253
Db		159	GlnGlulysPheAspAsp-----LeulleGlnSep-----	168
QY		254	GCGCCCTTTATCGAATTGCGGAACACAGCCTCGAATTTGCCAATACTACATTGGCGCAAC	313
Db		169	-----ileArgASnaIathrAspyfryileHisPheGIntyTyrrillelleGlnAsn	185
QY		314	GACATTTCCGGCAGCGTCTGTCCAACCTCATGTACCTTGGCGGAGAAGCGGGGTGGCG	373
Db		186	AspGluLeuGlyArgThrIleLeuAsnGluLeuGlyLysAlaGluGlnGlyValGlu	205
QY		374	GTAGCCTGTGTTGGAGCAGCAAACAACACGGCGGGTTGGACGATCTCCTGCTCGCCCTC	433
Db		206	VallysileLeutyAspMetGlySerargGlyLeu-----	218
QY		434	GACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCCTACGCAAAATGGCGC	493
Db		219	-----ArgLyLeSgLyLeuArGPProPheArgASNlysGlycylHis	232
QY		494	GCCTCGGTACCTGACCGACTTCCC-----CGCTCTCAAC	529
Db		233	AlaGluAlaPhe-----PhePrSerlYSLeuProLeuilleAsnLeuarGWmetAsn	249
QY		530	CGCGCATGCACAAACAATCTTTACCGCGCAACAACCGCGCACCATCTCGGGGGAGCG	589
Db		250	AsnArgAsnHISArgGLysIlleValValIleAspGlyGlnIleGlyTyrrValglycylPhe	269
QY		590	AATATCGCGCAGCAATACTTCAAAGTCGGTAGGACACCGGTTTTTCGCC-----GAC	640
Db		270	AsnValGlyaspGIutfr-----LeuGlyLYSSerlYSPhesGlyTyrrTrpArgasp	287
QY		641	CTGACATCTCTCGCCACCGCAGCGTCTGTCGGCGAAGTATCGCACGACTTCGACGGCTAC	700
Db		288	ThrHisleuArgIlleValGlyAspAlaValasnlalaLeuGlnLeuarGPheIleleuASP	307
QY		701	TGGCAAGCAATTCGCCGCCACAACGCCACGGCG-----	733
Db		308	TrpAsnSer-----GlnAlatHrarGasphIsIleSerTyrrAspAsArgtyrr	323

QY	734	-----ATCATCCGACGCGCAACATCGGCAAGGGTCTTCAAGACATCGGATCAAC	784
Db	324	PheProaspValasnSerGlyGlyThrIleGlyValGlnIleAlaSerSerGlyProasp	343
QY	785	GACGAACATCCAGACACGCGCTCTGCGCTACCGGAAACCGTCGAACAGTCGCCCCCTC	844
Db	344	Glulu	345
QY	845	TACAAAAAATACAGACGGGACGCATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGC	904
Db	346	TrpGlnIleIleTyserGlyTyserLeu	358
QY	905	GACACCCCTCGAAAAGGACTCGACCGCGCGCGCAACCCGCGATTCGCCGGAGGCTG	964
Db	359	SerAla	360
QY	965	CAAGACGCGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTCACTTATTCGTCCTC	1024
Db	361	LysLysSerIleTyserIleGlnSerProTyserPheIlePro	373
QY	1025	ACAAATCCGGCACAGACGCATCGCAAACTGGTCGACGAGCGGATAGACGTTACCGTC	1084
Db	374	AspGlnAlaPheLeuaspSerIleLysIleAlaAlaLeuGlyGlyValaspValasnIle	393
QY	1085	CTGACCAACTCGCTACAGCGCACCGCGTTCGCGCGTCCATTCGCGCTACGTCAAATAC	1144
Db	394	MetIleProasn	411
QY	1145	CGAAACCGGTCTCAAAACCGGCATCAAACTCTACGAGCTGCACCAACCATGCGGTC	1204
Db	412	AlaAlaSerLeuLeuaspAlaGlyValLysvalPheHisTyr	425
QY	1205	CCGCCACAAAAGACAAAGCGCTGACCGGCGCTCCGTAAACAGCGCTGCATGCCAAAACC	1264
Db	426	AspAsnGlyPhe	434
QY	1265	TTCATTGTGACGGCAAAACGCGATCTTCATCGGCTCATTCAACTCGACCCCGCTTCGCGCA	1324
Db	435	LeuValIleAspAspGluIleAlaSerValGlyThrAlaAsnMetAspHisargSerPhe	454
QY	1325	CGGCTCAATACCAATGGCGCTGTCATCGAAGCCCCCAAAATCCGACAAACAGATGGAG	1384
Db	455	ThrLeuAsnPheGluValAsnAlaPheIleTyserAspGlnGlnIleAlaLysLysLeuLys	474
QY	1385	CGCACCTCGCCGAT	1420
Db	475	GlnAlaPheIleaspAspLeuAlaValSerSerGluLeuThrLysAlaargTyrAlaLys	494
QY	1421	CGC 1423	
Db	495	Arg 495	
RESULT 32			
ABO53053			
ID	ABO53053 standard; protein; 956 AA.		
AC	ABO53053;		
XX	10-OCT-2003 (first entry)		
DT	Human putative spliceosome associated protein (SAP) #29.		
XX	Human; SAP; spliceosome associated protein; ribonucleoprotein;		
KW	RNP complex; RNA affinity substrate; RNP assembly sequence;		
KW	spliceosomal complex; hnRNP complex; mRNA export complex;		
KW	mRNA localisation complex; RNA editing complex; intron complex;		
KW	H complex; telomerase complex; fragile X protein complex;		
KW	reverse transcriptase complex; gene splicing complex.		
XX	Homo sapiens.		
OS	US2003068803-A1.		
XX			
PN			
XX			

Db	373	SerProGluLysLeuProGlnSerSerSerGluSerProProSerPro	: : : :	391
QY	250	CGCGCGCCCTTATCGAATCTCCGAACACAGCCTCGATTTCGAATACATACATTTGCGCG	: : : : : :	309
Db	392	ProThrLysValSerArgHisAlaSerSerSerProGluSerProAlaProAla	: : : : : :	411
QY	310	CAACGACATTTCGGCAGCGCTGCTGTTCAACCT	: : : : : :	342
Db	412	ProGlySerHisArgGluIleSerSerSerProThrSerLysAsnArgSerHisGlyArg	: : : : : :	431
QY	343	-----CATGTACCTTGCAGCAACCGCGCGCTGCGCGTACG	: : : : : :	378
Db	432	AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer	: : : : : :	451
QY	379	CTGTCTGTTGGACGACAAACACCGCGGTTGGACGATCTCT	: : : : : :	423
Db	452	ProAlaThrAlaLysArgGlyArgSerArg- --SerArgThrProThrLysArgGlyHis	: : : : : :	470
QY	424	GCTCGCCCTCGACAGCATCCCAATATCGAAGTGCCTGTTCAACCCCTTCGTCCTACG	: : : : : :	483
Db	471	SerArgSerArgSerProGlnTrpArgSerArgSerAlaGln	: : : : : :	485
QY	484	CAAAATGCGCGCACTCGGTACTGACCGACTTCCCGCGCTCAACCGCGCATGCACAA	: : : : : :	543
Db	486	ArgTyrGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro	: : : : : :	502
QY	544	CAAAATCCTTTACCGCGACACACCGCGCCACCATTCTCGCGGAGCGCAATATCGCGACGA	: : : : : :	603
Db	503	-----GlnArgProGlyTyrPsrArgSerArgAsnThrGlnArgArgGlyArg	: : : : : :	518
QY	604	ATACTTCAAAGTCGGTGAGGACCGCTTTTCGCGGACCTGCACATCTCGCCACCGCGCAG	: : : : : :	663
Db	519	SerArgSerAlaArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer	: : : : : :	538
QY	664	CGTCGTGCGGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCAATATCGCGCCACAA	: : : : : :	723
Db	539	ArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAlaArgArgArg	: : : : : :	558
QY	724	CGCACGCGCATCATCGGACGCGCAACAT	: : : : : :	762
Db	559	SerArgSerArgThrProThrArgArgArgSerArgSerArgThrProAlaArgArgGly	: : : : : :	578
QY	763	TCCTTAAGCACTCGGATA---CAACGACGAAACATCCAGACGCGCTCTCGCTACCG	: : : : : :	819
Db	579	ArgSerArgSerArgThrProAlaArgArgArgSerArgThrArgSerPro---ValArg	: : : : : :	597
QY	820	CGAAACGGTCGACAGTGCCTCTACCAAAAATACACACGGGCGCATCGCACTGCCA	: : : : : :	879
Db	598	ArgArgSerArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro	: : : : : :	617
QY	880	GAGCGTCCAAACCGCGCTATCAGCAGCACCCCTGCAAA-----AGACTCGACCGCA	: : : : : :	933
Db	618	AlaArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg	: : : : : :	637
QY	934	CGCGCGAAACCGCGATTCCCGGAGGCTGCAAGA-----	: : : : : :	969
Db	638	ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer	: : : : : :	657
QY	970	CGCGCTCAACACCGCGAAGGCGTCTATCTGTTTTCACCCCTATTTCGTCCTACAA	: : : : : :	1029
Db	658	ArgSerArgThrProArgGlyArgSerArgSerArgSerArgSerLeuValArgArgGlyArg	: : : : : :	677
QY	1030	ATCCGGACAGA-----CGCACTGCGAAACCTGGTCGAGCAGCGCATAGACGTTAC	: : : : : :	1080
Db	678	SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerGluArgLys---	: : : : : :	696
QY	1081	CGTCTCGACCACTCGCTACAGGCGACCGAGTTGCGC-----	: : : : : :	1119
Db	697	-----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet	: : : : : :	714
QY	1120	-----CGTCCATTCCGCTACGTCAAAATACCGAAACCGCTGCT	: : : : : :	1156

Db 715 LysLysSerArgIleSerSerArgSerArgSerLeuSerSerProArgSerLysala 734
 QY 1159 CAAGCCGGCATCAAACTCTACGA-----GCTGCAACCCCAACATGCGCTCCCGCCAC 1212
 Db 735 LysSerArgLeuSerLeuArgSerLeuSerGlySerSerProCys---ProLysGln 753
 QY 1213 AAAGACAAAGCGCTGACCGGAGCTCCGT-----AACAGCCTGC----- 1253
 Db 754 LysSerGlnThrProProArgArgSerArgSerGlySerGlnProLysAlaLysSer 773
 QY 1254 -----ATGCCAAACCTTCAATGTTGACGGCA 1280
 Db 774 ArgThrProProArgArgSerSerSerSerProProLysGlnLysSerLys 793
 QY 1281 AACGCATCTTCATCGGCTCATTCAACCTCGACCCCGCTCCG----- 1322
 Db 794 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 813
 QY 1323 -----CAGCGTCAATACCGAATGGCGTGCATCGAAGCCCAAAATCGCAG 1373
 Db 814 ProProArgGlnGlySerIleThrSerProGlnAlaSerGlnSerValThrProGln 833
 QY 1374 AACAGATGAGCGACCTCCCGCATACACACCGAATACGCTACCGCTTACCTCG 1433
 Db 834 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 852
 QY 1434 ACAACACCAACCGCTCGAATGGC-----ACGATCCCGCA 1469
 Db 853 ArgHis-----SerCysSerGlySerSerProProArgValLysSerSerThrProPro 870
 QY 1470 CCGAATAACTACCGCA---ACGAACCGAAGCA----- 1502
 Db 871 ArgGlnSerProSerArgSerSerSerProGlnProLysValLysAlaIleSerPro 890
 QY 1503 -----ACCTTGGAAACGATCGCGCAAAATCCTATCCTGCTGCCATCG 1550
 Db 891 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 904
 QY 1551 AAGGTT 1556
 Db 905 ArgVal 906
 RESULT 33
 ADN60280
 ID ADN60280 standard; protein; 2296 AA.
 XX
 AC ADN60280;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human splicing coactivator subunit SRM300 protein.
 XX
 KW angiogenesis modulator; tumorigenesis modulator; angiogenesis; Ax1;
 KW tubulin cofactor D; transglutaminase 2; cytosine deaminase;
 KW peptidase M41; paraplegin; CD13 aminopeptidase; PPK-1; zip kinase; Gas6;
 KW SRM160; non-muscle myosin heavy chain; calmodulin 2; symporter;
 KW semaphorin; zinc finger helicase; plexin-A2; deoxycytidylate deaminase;
 KW sugar transporter; tumorigenesis; antiangiogenic; cytosolic;
 KW cerebroprotective; vasotropic; antiinfertility; cardiant;
 KW antibody therapy; antisense therapy; RNA interference therapy;
 KW RNAi therapy; cancer; stroke; infertility; heart disease; human.
 XX
 OS Homo sapiens.
 XX
 PN W02004039955-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 29-OCT-2003; 2003WO-US034281.
 XX
 PR 29-OCT-2002; 2002US-0421989P.
 PR 17-OCT-2003; 2003US-0512251P.
 XX

(RICE-) RIGEL PHARM INC.

Lorens JB, Atchison RE, Frieria A, Holland S;

WPI; 2004-376181/35.

N-PSDB; ADN60279.

Identifying a compound that modulates angiogenesis or tumorigenesis, useful in diagnosing and treating angiogenesis, cancer, stroke, infertility and heart disease, comprises contacting the compound with angiogenesis polypeptide.

Disclosure; Page 70-71; 105pp; English.

The present invention describes a method for identifying a compound that modulates angiogenesis or tumorigenesis. The method comprises: (a) contacting the compound with angiogenesis polypeptide, e.g. Ax1, tubulin cofactor D, transglutaminase 2, cytosine deaminase, peptidase M41 (paraplegin), CD13 aminopeptidase, PPK-1, zip kinase, Gas6, SRM160, non-muscle myosin heavy chain, calmodulin 2, novel symporter, novel semaphorin, novel zinc finger helicase (FLJ22611), plexin-A2, deoxycytidylate deaminase or novel sugar transporter; (b) determining the functional effector of the compound upon the angiogenesis polypeptide or the physical effect of the compound upon the target polypeptide or its fragment or inactive variant; and (c) determining the chemical or phenotypic effect of the compound upon a cell comprising the target polypeptide or its fragment or inactive variant, thus identifying a compound that modulates cell cycle arrest. Also described is a method of modulating angiogenesis in a subject. The angiogenesis or tumorigenesis modulating compound has antiangiogenic, cytostatic, cerebroprotective, vasotropic, antiinfertility and cardiant activities, and can be used in antibody, antisense and RNA interference (RNAi) therapies. The method is useful in identifying a compound that modulates angiogenesis. The methods and compounds or compositions are useful in diagnosing and treating angiogenesis, cancer, stroke, infertility and heart disease. The present sequence represents a human splicing coactivator subunit SRM300 protein, which is used in the exemplification of the present invention.

Sequence 2296 AA;

Alignment Scores:

Pred. No.:	1,79e-13	Length:	2296
Score:	303.50	Matches:	156
Percent Similarity:	37.6%	Conservative:	78
Best Local Similarity:	25.1%	Mismatches:	233
Query Match:	10.6%	Indels:	156
DB:	8	Gaps:	29

US-10-665-990A-13 (1-1561) x ADN60280 (1-2296)

QY	43	AACACGAGCCTCATTTCCCTTTTATGCTCTCTCTCATGTCTTCATGTTGCC	102
Db	358	SerThrGlyProGluProProAlaProThrProLeuLeu-----	370
QY	103	CCCACTGGAAGAACGACGAGGAAAGCGCTCATTTCAATCTTCCAAACCTGCTCCTCTGGA	162
Db	371	---AlaGluArgHisGlyGlySerPro-----GlnProLeuAlaThrThrProLeuSer	387
QY	163	CAACATCTCTGCAAAATCCG-----GCACACCTCTATAACACGGGTATCCGACAT	213
Db	388	GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProLys	407
QY	214	CTACCTGCTCGACGACCC-----CCACGAAGCCTTGCGCG	249
Db	408	SerProGluLysLeuProGlnSerSerSerSerGluSerSerProProSerPro---Gln	426
QY	250	CCGCGCGCGCTTATCGAATCTGCCGAACACACACCTCGATTTCGAATACATACATTGGCG	309
Db	427	ProThrLysValSerArgHisAlaSerSerSerProGluSerProAlaProAla	446
QY	310	CAACGACATTTCCGAGGCTGCTGTCAACCT-----	342
Db	447	ProGlySerHisArgGluIleSerSerSerProThrSerLysAsnArgSerHisGlyArg	466

QY 343 -----CATGTACCTTCCCGCAGAACCGCGGTCCGGCTACG 378
 Db 467 AlalysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
 QY 379 CCTGCTGTGGACGACAAACACACCGCGGGTGGACGATCCCT- 423
 Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
 QY 424 GCTCCCTTCGACAGCCATCCCAATATCGAAGTCGCCTGTTCACCCCTTCGTCTACG 483
 Db 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerAlaGln----- 520
 QY 484 CAAATGGCGCGACTCGGTACTACCGACTTCCCGCCCTCAACCGCGCATCCACAA 543
 Db 521 ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro- 537
 QY 544 CAAATCCTTTACCGCGCAGAACCGCGCACCATCTACGCGGAGCAATATCGGCGAGA 603
 Db 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
 QY 604 ATACTTCAAAGTCGTGAGCACCCGTTTTCGCCGACCTGGACATCTCTCGCCACCGGAG 663
 Db 554 SerArgSerAlaArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
 QY 664 CGTCGTCGGGAGATATCGCAGACTTCGACCGCTACTGGGCAAGCCATTCCGCCACAA 723
 Db 574 ArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAlaArgArg 593
 QY 724 CGCCACGGCGCATCATCCGCGAGCGCAACAT-----CGGCAAGGG 762
 Db 594 SerArgSerArgThrProThrArgArgArgSerArgSerArgThrProAlaArgArgGly 613
 QY 763 TCTTCAAGCACTCGGATA---CAACGACGAAACATCCAGACAGCGCTCTCGGTACCG 819
 Db 614 ArgSerArgSerArgThrProAlaArgArgSerArgSerArgThrArgSerPro---ValArg 632
 QY 820 CGAACCCTCGAACAGTCCGCCCTCTACCAAAAATACAGACGGAGCGCATCGATGGCA 879
 Db 633 ArgArgSerArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro 652
 QY 880 GAGCGTCCAAACCCGCTGATCAGCGACACCCCTGCAAA-----AGGACTCGACCGCA 933
 Db 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
 QY 934 CCGCGCAAAACCGCGATTCGCGGAGGCTGCAAGA----- 969
 Db 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
 QY 970 CGCGCTCAACAGCCGGAAGCGTCTATCTGGTTTCACCTATTTCGTCCCTACAA 1029
 Db 693 ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg 712
 QY 1030 ATCCGCGCACAGA-----CGCACTGGCAAACTGGTGACGAGCGGATAGAGCTTAC 1080
 Db 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerSerGluArgLys--- 731
 QY 1081 COTCTGTACCAACTCGCTACAGCGGACCGAGCTGTCGCG- 1119
 Db 732 -----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet 749
 QY 1120 -----CGTCCATTCCGGTACGTCAATACCGAAACCGCTGCT 1158
 Db 750 LysLysSerArgIleSerSerArgArgSerArgSerLeuSerSerProArgSerLysAla 769
 QY 1159 CAAAGCGCGCATCAAACTCTACGA-----GTCGAACCAACCATCGCTCCCGCCAC 1212
 Db 770 LysSerArgLeuSerLeuArgArgSerLeuSerGlySerSerProCys---ProLysGln 788
 QY 1213 AAAAGACAAAGCGCTGACCGGAGCTCCGT-----AACGAGCTGC- 1253
 Db 789 LysSerGlnThrProProArgArgSerArgSerGlySerGlnProLysAlaLysSer 808

QY 1254 -----ATGCCAAACCTTCATTGTGGACGGCA 1280
 Db 809 ArgThrProProArgArgSerArgSerSerSerProProLysGlnLysSerLys 828
 QY 1281 AACGCATCTTCATCGGCTCATTCACCTCGACCCCGTTCG- 1322
 Db 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
 QY 1323 -----CAGCGCTCAATACCGAATGGCGGTGCTCATCGAAAGCCCCAAATCGCAG 1373
 Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluGlnSerValThrProGln 868
 QY 1374 AACAGATGAGCGCACCCCTCCCGCATACACACCGAATACCGCTACCGGTTACCTCG 1433
 Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
 QY 1434 ACAAAACAACCGCTGCAATGGC-----ACGATCCCGGCA 1469
 Db 888 ArgHis-----SerCysSerGlySerSerProProArgValLysSerSerThrProPro 905
 QY 1470 CCGGAAACCTACCCGA---ACGAACCCGGAAGCA- 1502
 Db 906 ArgGlnSerProSerArgSerSerProGlnProLysValLysAlaIleSerPro 925
 QY 1503 -----AACTTTGAAACGATCGCGCGCAAAAATCCTATCCCTGCTGCCCATCG 1550
 Db 926 ArgGlnArgSerHisSerGlySerSerSerProSer- 1551
 QY 1551 AAGGTT 1556
 Db 940 ArgVal 941
 RESULT 34
 ADN60278
 ID ADN60278 standard; protein; 2752 AA.
 XX AC ADN60278;
 XX DT 18-NOV-2004 (first entry)
 XX DE Human serine/arginine repetitive matrix 2 protein.
 KW angiogenesis modulator; tumorigenesis modulator; angiogenesis; Axl;
 KW tubulin cofactor D; transglutaminase 2; cytosine deaminase;
 KW peptidase M41; paraplegin; CD13 aminopeptidase; PKP-1; zip kinase; Gas6;
 KW SRM160; non-muscle myosin heavy chain; calmodulin 2; symporter;
 KW semaphorin; zinc finger helixase; plexin-A2; deoxycytidylate deaminase;
 KW sugar transporter; tumorigenesis; antiangiogenic; cytosolic;
 KW cerebrotective; vasotropic; antifertility; cardiac;
 KW antibody therapy; antisense therapy; RNA interference therapy;
 KW RNAi therapy; cancer; stroke; infertility; heart disease; human.
 XX OS Homo sapiens.
 XX WO2004039955-A2.
 XX PD 13-MAY-2004.
 XX PF 29-OCT-2003; 2003WO-US034281.
 XX PR 29-OCT-2002; 2002US-0421989P.
 XX PR 17-OCT-2003; 2003US-0512251P.
 XX XX (RIGEL-) RIGEL PHARM INC.
 XX XX Lorens JB, Atchison RE, Frier A, Holland S;
 XX WPI: 2004-376181/35.
 XX DR N-PSDB; ADN60277.
 XX PT Identifying a compound that modulates angiogenesis or tumorigenesis,
 PT useful in diagnosing and treating angiogenesis, cancer, stroke,
 PT infertility and heart disease, comprises contacting the compound with

angiogenesis polypeptide.

Disclosure; Page 68; 105pp; English.

The present invention describes a method for identifying a compound that modulates angiogenesis or tumorigenesis. The method comprises: (a) contacting the compound with angiogenesis polypeptide, e.g. Axl, tubulin cofactor D, transglutaminase 2, cytosine deaminase, peptidase M41 (paraplegin), CD13 aminopeptidase, PK-1, zip kinase, Gas6, SRm160, non-muscle myosin heavy chain, calmodulin 2, novel symporter, novel semaphorin, novel zinc finger helicase (FLJ22611), plexin-A2, deoxycytidylate deaminase or novel sugar transporter; (b) determining the functional effector of the compound upon the angiogenesis polypeptide or the physical effect of the compound upon the target polypeptide or its fragment or inactive variant; and (c) determining the chemical or phenotypic effect of the compound upon a cell comprising the target polypeptide or its fragment or inactive variant, thus identifying a compound that modulates cell cycle arrest. Also described is a method of modulating angiogenesis in a subject. The angiogenesis or tumorigenesis modulating compound has antiangiogenic, cytostatic, cerebroprotective, vasotropic, antiinfertility and cardiact activities, and can be used in antibody, antisense and RNA interference (RNAi) therapies. The method is useful in identifying a compound that modulates angiogenesis. The methods and compounds or compositions are useful in diagnosing and treating angiogenesis, cancer, stroke, infertility and heart disease. The present sequence represents a human serine/arginine repetitive matrix 2 protein, which is used in the exemplification of the present invention.

Sequence 2752 AA;

Alignment Scores:

Pred. No.: 1.87e-13 Length: 2752
Score: 303.50 Matches: 156
Percent Similarity: 37.6% Conservative: 78
Best Local Similarity: 25.1% Mismatches: 233
Query Match: 10.6% Indels: 156
DB: 8 Gaps: 29

US-10-665-990A-13 (1-1561) x ADN60278 (1-2752)

QY 43 AACACGAGCTCATTTCCCTTTATGCTCTCTGTTTCATGTTCTTCATGTTGCC 102
DB 358 SerThrGlyProGluProAlaProThrProLeuLeu----- 370
QY 103 CCCACTGGAAGACGAGCGAAAGCGGTCATTTCATTAATCTCCAAACCTGTCCTCTGGA 162
DB 371 ---AlaGluAArgHisGlySerPro-----GlnProLeuAlaThrProLeuSer 387
QY 163 CAACATCTGCAATTCG-----GCACACCCCTCATAAACCGGCTATCCGACAT 213
DB 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProlys 407
QY 214 CTACTGCTCAGACACC-----CCAGAGACCTTGGCGC 249
DB 408 SerProGluLysLeuProGlnSerSerSerSerGluSerSerProSerPro---Gln 426
QY 250 CCGCGCCGCTTATCGAATCTGCGAACACAGCTCGATTGCAATACATTTGGCG 309
DB 427 ProThrLysValSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446
QY 310 CAACGACATTTCCGGCAGGCTGTTCAACCT----- 342
DB 447 ProGlySerHisArgGluLeuSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
QY 343 -----CATGTACTTGGCGCAGAACCGCGGCTGCGGTACG 378
DB 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
QY 379 CCTGCTGTGGACGACACACACGCGGGTGGACGATCTCCT----- 423
DB 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
QY 424 GCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCGTCCTACG 483

DB 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerAlaGln----- 520
QY 484 CAATATGGCGCATCTCGGTACTTGCACGACTTCCCGCGCTCAACCCGCGCATGCACAA 543
DB 521 ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro----- 537
QY 544 CAATATCTTTACCCTCCGACAAACCGCGCCACCATATCTCGCGGAGCAGCAATATCGCGCAGA 603
DB 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
QY 604 ATACTTCAAACCTCGGTGAGACACCGTTTTCGCGACCTCGACATCTCTCGCCACCGCGAG 663
DB 554 SerArgSerAlaArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
QY 664 CGTCGTGGCGCAAGTATCGCAGACTTCGACCGTCTGCGCAAGCCATTCGCGCCACAA 723
DB 574 ArgSerArgThrProAlaArgGlyArgSerArgSerArgSerArgThrProAlaArgArg 593
QY 724 CGCCACGCGCATCATCCGACGCGCACAT-----CGGCAAGG 762
DB 594 SerArgSerArgThrProThrArgArgArgSerArgSerArgThrProAlaArgArgGly 613
QY 763 TCTTTCAAGCACTCGGATA---CAACGACGAAACATCCAGACACGCGCTCTCTCGGTACCG 819
DB 614 ArgSerArgSerArgThrProAlaArgArgSerArgThrArgSerPro---ValArg 632
QY 820 CGAAACCGTGAACAGTCGCCCCCTCTACCAAAAATACAGACGGGACGCATCGACTGCGA 879
DB 633 ArgArgSerArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro 652
QY 880 GAGCGTCCAAACCGCGCTGATCAGCGACACCCCTGCAAA-----AGGACTCGACCGGA 933
DB 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
QY 934 CCGCGCAAAACCGCGGATTGCGGGAGGCTGCAAGA----- 969
DB 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
QY 970 CGCGCTCAACAGCGCCGAAACACGCTCTATCTGTTTCCACCTATTTCGTCCCTACAAA 1029
DB 693 ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg 712
QY 1030 ATCCGCGCACAGA-----CGCACTGGCAAACTGTGCGAGCAGCGCATAGAGTTAC 1080
DB 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerSerGluArgLys--- 731
QY 1081 CGTCTGACCAACTCGCTACAGCGCGACGCTTTCGCGC----- 1119
DB 732 -----AsnLysSerArgThrSerGlnArgSerArgSerArgSerAsnSerSerProGluMet 749
QY 1120 -----CGTCCATTCCGGCTAGCTCAATACCGAATAACCGAACCCTGCT 1158
DB 750 LysLysSerArgLysSerArgSerArgSerArgSerLeuSerSerProArgSerLysala 769
QY 1159 CAAGCGCGCATCAAACTCTACGA-----GCTCAACCCCAACCATCGCGTCCCGCCAC 1212
DB 770 LysSerArgLeuSerLeuArgSerArgSerLeuSerLysSerSerProLysGln 788
QY 1213 AAAAGACAAAGGCTGACCGCGCAGCTCCGT-----AACCAAGCTGC----- 1253
DB 789 LysSerGlnThrProProArgArgSerArgSerGlySerGlnProLysAlaLysSer 808
QY 1254 -----ATGCCAAAACCTTCATTCTGTGGACGCGCA 1280
DB 809 ArgThrProProArgArgSerArgSerSerSerSerProProLysGlnLysSerLys 828
QY 1281 AACGCATCTTCATCGGCTCATCAACCTCGACCCCGCTCCG----- 1322
DB 829 ThrProSerArgGlnSerHisSerSerSerSerProHisProLysValLysSerGlyThr 848
QY 1323 -----CACGGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAG 1373
DB 1323 -----CACGGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAG 1373

Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluSerValThrProGln 868
QY 1374 AACAGATGAGCGCACCTCGCGGATACACACCGAATACCGCTTACCGCTCG 1433
Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
QY 1434 ACAACACACACCGCTCGCATGGC-----ACGATCCCGGCA 1469
Db 888 ArgHis-----SerCysSerGlySerSerProGlnProValLysSerSerThrProPro 905
QY 1470 CCGGAAAAACCTACCCGA---ACGAACCGGAAGCCA----- 1502
Db 906 ArgGlnSerProSerArgSerSerProGlnProLysValLysAlaIleIleSerPro 925
QY 1503 -----AAGTTTGGAAACGATCGCGCGCAAAAATCCTATCCCTGCTGCCCATCG 1550
Db 926 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 939
QY 1551 AAGGTT 1556
Db 940 ArgVal 941
RESULT 35
ADP24102
ID ADP24102 standard; protein; 2752 AA.
XX
AC ADP24102;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide SEQ ID NO:1280.
XX
KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu ID;
XX
DR WPI; 2004-419628/39.
DR N-PSDB; ADP24101.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 1280; 2940pp; English.
XX

CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune

CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX

SEQ Sequence 2752 AA;

Alignment Scores:
Pred. No.: 1.87e-13 Length: 2752
Score: 303.50 Matches: 156
Percent Similarity: 37.6% Conservative: 78
Best Local Similarity: 25.1% Mismatches: 233
Query Match: 10.6% Indels: 156
DB: 8 Gaps: 29

US-10-665-990A-13 (1-1561) x ADP24102 (1-2752)

QY 43 AACACGACGCTCATTTCCCTTTTATGCTCTCCTCTCTTTCATGTTCTTCTCATGTTGCC 102
Db 358 SerThrGlyProGluProProAlaProThrProLeuLeu----- 370
QY 103 CCCACTGGAAGAACGGACGGAAGCGTCATTTCAATCTCTCCAAACCTGCTCTCTGGA 162
Db 371 ---AlaGluArgHisGlyGlySerPro-----GlnProLeuAlaThrThrProLeuSer 387
QY 163 CAACATCTCGCAATCCG-----GCACACCCCTCATACACAGCGCTATCCGACAT 213
Db 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProLys 407
QY 214 CTACCTGCTCGACGACCC-----CCACGAAGCCCTTGCCGC 249
Db 408 SerProGluLysLeuProGlnSerSerSerGluSerSerProProSerPro---Gln 426
QY 250 CCGCGCCGCTTATCGAATCTGCGAACACAGCTCGATTGTCATCTACTACTATTGGCG 309
Db 427 ProThrLysValSerArgHisAlaSerSerProGluSerProLysProAlaProAla 446
QY 310 CAACGACATTTCCGCGAGCTGCTGTTCAACCT----- 342
Db 447 ProGlySerHisArgGluLeuSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
QY 343 -----CATGTACCTTCCGAGAACCGCGCTGCGGTACG 378
Db 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
QY 379 CTGCTGTGTGGACGACACACACCGCGGTGGACGATCTCCT----- 423
Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
QY 424 GTCGCGCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCTGCTTACG 483
Db 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerArgAlaGln----- 520
QY 484 CAAATGGCGCGCACTCGGCTACTCGACCGACTTCCCGCCCTCAACCGCGGATGCACAA 543
Db 521 ArgTpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro----- 537
QY 544 CAAATCTTTACCGCGCAACCGCGCCACCATCTCGCGGCGGACCAATATCGGCGCGCA 603
Db 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
QY 604 ATACTTCAAAGTCGGTGAGGACACCGTTTTTCGCCGACCTGGACATCCTCGCCACCGCGAG 663

554	Ser	Arg	Ser	Ala	Arg	Arg	Gly	Arg	Ser	His	Ser	Arg	Ser	Pro	Ala	Thr	Arg	Gly	Arg	Ser	573
664	CGT	CGT	CGG	CAAG	TAT	CGC	ACG	ACTT	CGC	ACG	CGT	TACT	TGG	CAAG	CAAG	CACTT	CGG	CCCA	CA	723	
574	Arg	Ser	Arg	Thr	Pro	Ala	Arg	Arg	Gly	Arg	Ser	Arg	Ser	Arg	Thr	Pro	Ala	Arg	Arg	593	
724	CGC	CAC	CGG	CAT	ATC	CGC	ACG	CGCA	CAAT	-----	-----	-----	-----	CGC	AA	GGG	762				
594	Ser	Arg	Ser	Arg	Thr	Pro	Thr	Arg	Arg	Arg	Ser	Arg	Ser	Arg	Thr	Pro	Ala	Arg	Arg	613	
763	TCTT	CAAG	CACT	TCG	GAT	---CA	ACG	AGAA	CAAT	CC	CAG	ACAC	CGG	CTCT	CGC	CTAC	CG	819			
614	Arg	Ser	Arg	Ser	Arg	Thr	Pro	Ala	Arg	Arg	Ser	Arg	Thr	Arg	Thr	Arg	Ser	Pro	---Val	632	
820	CGAA	ACCG	TG	CAAC	AGT	CGC	CCCT	CTT	AC	CAAA	AAAT	AT	CAG	ACG	CA	CGC	AT	CG	CA	879	
633	Arg	Arg	Ser	Arg	Ser	Arg	Ser	Pro	Ala	Arg	Ser	Gly	Arg	Ser	Arg	Ser	Arg	Thr	Pro	652	
880	GAG	CGT	CAAA	CCG	CCG	TG	AT	CAG	CA	CC	CT	GC	AAA	-----	AG	CT	CA	CG	CG	933	
653	Ala	Arg	Gly	Arg	Ser	Arg	Thr	Pro	Ala	Arg	Arg	Gly	Arg	Ser	Arg	Ser	Arg	672			
934	CCG	CG	CAAA	CCG	CG	AT	TG	CGG	AGG	CT	CA	AG	-----	969							
673	Thr	Pro	Ala	Arg	Arg	Ser	Gly	Arg	Ser	Arg	Ser	Arg	Thr	Pro	Ala	Arg	Gly	Arg	Ser	692	
970	CGC	GT	CAAA	CAG	CG	CAAA	AAAG	CG	CT	AT	CT	CGT	TTC	AC	CC	TAT	TTC	CG	CC	1029	
693	Arg	Ser	Arg	Thr	Pro	Arg	Gly	Arg	Ser	Arg	Ser	Arg	Ser	Leu	Val	Arg	Gly	Arg	712		
1030	ATC	CGG	CAC	AGA	-----	CG	CAT	TG	CG	CAAA	AACT	GT	CG	CAG	CG	CA	TAG	AG	CT	1080	
713	Ser	His	Ser	Arg	Thr	Pro	Gln	Arg	Arg	Gly	Arg	Ser	Gly	Ser	Ser	Glu	Arg	Lys	---731		
1081	CGT	CCT	GAC	CAACT	CG	CT	AC	AGG	CG	ACG	AGT	TG	CGC	-----	1119						
732	-----	Asn	Lys	Ser	Arg	Thr	Ser	Gln	Arg	Arg	Ser	Arg	Ser	Asn	Ser	Ser	Pro	Glu	Met	749	
1120	-----	-----	CGT	CCAT	TCC	GGT	TAC	GT	CAAT	AC	CAAA	AC	CG	CT	GT	1158					
750	Lys	Lys	Ser	Arg	Gln	Ser	Arg	Ser	Arg	Ser	Leu	Ser	Leu	Ser	Pro	Arg	Ser	Lys	Ala	769	
1159	CAA	AGC	CGG	CACT	CAAA	CT	CT	AC	GA	-----	GCT	CAAC	CCCA	ACCAT	CGC	CT	CCCG	CG	CA	1212	
770	Lys	Ser	Arg	Leu	Ser	Leu	Arg	Ser	Leu	Ser	Ser	Gly	Ser	Pro	Cys	---Pro	Lys	Gln	788		
1213	AAA	AG	CAAA	AGG	CT	AC	CGG	CACT	CGT	-----	---AAC	AGC	CT	CG	-----	1253					
789	Lys	Ser	Gln	Thr	Pro	Pro	Arg	Arg	Ser	Ser	Gly	Ser	Gly	Ser	Gln	Pro	Lys	Ala	Lys	808	
1254	-----	-----	-----	ATG	CAAA	ACCT	CT	TAT	TGT	TG	CA	CG	CA	1280							
809	Arg	Thr	Pro	Pro	Arg	Arg	Ser	Arg	Ser	Ser	Ser	Pro	Pro	Pro	Lys	Gln	Lys	Ser	Lys	828	
1281	AAC	GCAT	TCT	TAT	CGG	CT	CAAT	CAAC	CT	CGAC	CC	CCG	TT	CCG	-----	1322					
829	Thr	Pro	Ser	Arg	Gln	Ser	His	Ser	Ser	Ser	Pro	His	Pro	Lys	Val	Lys	Ser	Gly	Thr	848	
1323	-----	-----	CAC	GGT	CAAT	AC	CGA	AA	TGGG	CT	CGT	CAT	CGA	AA	CGCC	CAAA	AT	CG	CAG	1373	
849	Pro	Pro	Arg	Gln	Gly	Ser	Ile	Thr	Ser	Pro	Gln	Ala									


```
QY 1423 -----CGTTACCTCGACAAACACACCGCTGCAATGCGCAGA 1461
Db 579 ArgTrpArgAlaAspHisArgHisProArgGlnAlaThrAspSerArgHisAlaHis 598
QY 1462 TCCCGCCGCCCAAAACCTACCCGAGAACCGACCGAAGCAAACTTTGGAAACGCATCGC 1521
Db 599 GlyArgHisProArgAsp-----Arg 605

QY 1522 CCACAA-----AATCTATCCCTGCTGCCCATCGA 1551
Db 606 ArgArgAlaGlyGlyArgArgLeuProAlaValHisArg 618

RESULT 38
ABO74950
ID ABO74950 standard; protein; 618 AA.
XX
AC ABO74950;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #7125.
XX
DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD08521.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 23696; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 618 AA;

Alignment Scores:
Pred. No.: 1.68e-13 Length: 618
Score: 302.00 Matches: 173
Percent Similarity: 34.8% Conservative: 35
Best Local Similarity: 29.0% Mismatches: 193
Query Match: 10.6% Indels: 196
```

```
DB: 7 Gaps: 38
US-10-665-990A-13 (1-1561) x ABO74950 (1-618)
QY 169 CCTGCAAAATCGGCACACCCCTCATACAAACAGGGCTATCCGACATCTACCTGCTCGACGA 228
Db 4 ProAspHisProArgHisProArg---ArgArgGlnLeuAlaGluLeuArgIleArgArg 22
QY 229 -----CCCCACGAAGCCCTTTGCGCGCCGCGCCCTTATCGAATCTGCGCAACA 279
Db 23 ArgThrAlaProGlySerAspArgGlnGlnProArgArgThrProArgIleArgArgArg 42
QY 280 CAGCCTCGATTGCAATACTACATTGTCGCGCAACGACAT----- 318
Db 43 -----HisGlnGlyGlnSerHisArgProAlaHisGlnGlyArg 55
QY 319 -----TTCCGGCAGGCTGCTTCAACCTCAT 345
Db 56 GlnArgGlnProGlyAlaProAlaAlaMetGlyLeuArgArgAla-----TrpProAla 73
QY 346 GTACCTTCCCGCAGAACCGCGTGCCTGCTGTGGAGCAGACAAACACCGG 405
Db 74 AlaProCysGlyArgArgArgProAspThr-----GlnLeuArgLeu 88
QY 406 CGGGTTGACGATCTCTGCTGCGCCCTCGACAGCCATCCCAATATCGAAGTCGCGCTGTT 465
Db 89 ArgProGlnArgGlnSerGlyArgArgAsnGlnPro-----AlaProVal 103
QY 466 CAACCCCTT-----ATCCTTTACCGCGCAACCG----- 567
Db 144 ArgCysHisHisProLeuArgIleArgArgProArgGlnProAspProThrGlyGlnPro 163
QY 568 -----CGCCACCACTACTCGCGGACGCAATATCGCGACGAATACTTCAAAAGTCGGTGA 621
Db 184 GlyGlnArgHisHisHisLeuArgAlaArgArgArg-----GlnArgHisPro 180
QY 622 GGACACCGTTTTGCGCCGACCTGACATCTCGCCACCGCGACGCTGTCGGCGAAGTATC 681
Db 181 ProHisArgArgSerArgCysGlyHis---ArgValSerLeuArgArgProGlnSerAla 199
QY 682 GCACGACTTCGACCGCTACTGGGCAAGCATTCGCCCCACACGCGCATCATCCG 741
Db 200 GlyArg---AlaProLeuAlaGluArgPro---GluProArgArgThrVal-----Pro 215
QY 742 CAGCGGCAACATCGGCAAGGCTTCAAGCACTCGGATACAA---CGACGAAACATCCAG 798
Db 216 LeuArgProHisArgArgArgGlnGlnGlyHisArgProProGlyArgHisArgTrpArg 235
QY 799 ACACGCGCTCTCGCTACCGCGAAACCGTCGAAACAGTCGCCCTCTACCAAAAAATACA 858
Db 236 SerArgGlnProGlyValProLeuArgArg----- 245
QY 859 GACGGGCGCATCGACTGGCAGCGTCCAAACCCGCTGATCAGCAGACACCCCTGCAAA 918
Db 246 -----AlaArgGlnProGlyArgAlaGlyAlaGlnHisPro----- 257
QY 919 AGGACTCGACCGCGACCGCCGCAACCGCATTCGCGGAG---GCTGCAAGACGCGCT 975
Db 258 -----ProArgProAlaAspProAlaArgProArgAspLeuProLeuArgArgGly 274
QY 976 CAACACGCGCGAAAAAAGCGTCTATCTGGTTTACCCCTATTTCGCTTACAAAATCCGG 1035
Db 275 GlnProThrAlaGlyAspArgLeu-----ProLeuArgProArg 287
QY 1036 CAC-----AGACGCACTGCCAA-----ACTGGTCGACGACCG 1068
```

288	HisArgLeuProAlaGlnArgArgProGlyArgGlnArgAspProGlySerGlyArg	307
1069	CATAGACGTTACCGTCTCT-----GACCAACTCGCTACA-----	1101
308	GlnGlyAlaGluHisProGlyArgAlaAspArgLeuProAlaLeuArgProAlaAlaAa	327
1102	-----GGCGACGACGTTGGCGGTCCATCCGGCTACGTCAA---ATACCGAAA	1149
328	ProAspLeuGlnArgHisSerGlnPro---ArgValArgProGlyLeuProThr	346
1150	ACCGCTGCTCAAAGCCGG-----CATCAAACTCTACGAGCTGCAACCCCAA-	1194
347	AlaAlaAlaGluArgProLeuAlaGluArgLeuProAlaArgCysGlnTrpGlnTyr	366
1195	-----CGATGCCGTCCCCGCCACAAA	1215
367	ProAlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro	386
1216	AGACAAAGCGCTGACCGCGAGCTCGGTAAACGAGCGCTCATGCCAA---AACCTTCATTGT	1272
387	-----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuCysLeuArgArg	403
1273	GGACGCCAAACGCAT-----CTTCATCGCGTCAATTCAA	1305
404	GlyArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---Gln	422
1306	CCTCGACCCCGTTCCGACGGCTCAATAC-----CGAAATGGCGTGTGTCATCGA	1356
423	ProGlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgAlaSerGly	442
1357	AAGCCCAAATCCAGAACAGATGGAGCGCACCTCGCGCATACCACCCCGAATACGC	1416
443	AspGlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArg	462
1417	CTACCGGTTACCTCGACAAACACACCGCTCGCAATGGCA-----	1458
463	ArgProGlyProProAlaLysArgGlnProArgArgProAlaGlyArgArgIleProLeu	482
1459	-----CGATCCGGCCACCCGAAA-----AACCTTACCC---	1485
483	GlnArgProArgProAlaHisArgGlnAlaHisProArgGluArgHisLeuProLeu	502
1486	-----GAACGACCCCGAAGCCAAACTTTGGAAACGCAT	1518
503	TrpSerArgProValAlaArgArgSerArgThrArgArgGlnArgProGluThrAla	522
1519	CGCCGCAAAATCTATC-----CTGTGTGCC-----CATCGAAGG	1554
523	SerProValLeuProValAlaGlnProAlaAlaGlyHisHisArgArg	539

RESULT 39	
ADJ70425	
ID	ADJ70425 standard; protein; 2263 AA.
XX	
AC	ADJ70425;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Human heat mitochondrial protein as a therapeutic target SeqID2231.
XX	
DE	mitochondrial; human; screening assay; diabetes mellitus;
XX	
KW	Huntington's disease; osteoarthritis;
KW	Leber's hereditary optic neuropathy; LHON;
KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW	neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW	osteopathic; ophthalmological; cytostatic.
XX	
OS	Homo sapiens.
XX	
PN	WO2003087768-A2.
XX	

PD	23-OCT-2003.	
XX		
PF	04-APR-2003; 2003WO-US010870.	
XX		
PR	12-APR-2002; 2002US-0372843P.	
PR	17-JUN-2002; 2002US-0389987P.	
PR	20-SEP-2002; 2002US-0412418P.	
XX		
PA	(MITO-) MITOKOR.	
PA	(BUCK-) BUCK INST AGE RES.	
XX		
PI	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;	
PI	Warnock DE;	
XX		
DR	WPI; 2003-845369/78.	
XX		
PT	Identifying a mitochondrial target for drug screening assays and for	
PT	treating diseases associated with altered mitochondrial function,	
PT	comprises detecting a modified polypeptide in a sample and correlating	
PT	with the disease.	
XX		
PS	Claim 1; SEQ ID NO 2231; 180pp; English.	
XX		
CC	This invention relates to novel mitochondrial targets that can be used	
CC	for therapeutic intervention in treating a disease associated with	
CC	altered mitochondrial function. Specifically, it refers to a method for	
CC	identifying proteins of the human heart mitochondrial proteome that are	
CC	useful for drug screening assays, as well as therapeutic targets. The	
CC	present invention describes a method for identifying such proteins that	
CC	can be used in the treatment of various diseases associated with altered	
CC	mitochondrial function including diabetes mellitus, Huntington's disease,	
CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial	
CC	encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy	
CC	raged red fibre syndrome (MERRF) or cancer. Accordingly, these	
CC	compositions have neuroprotective, nootropic, antidiabetic,	
CC	anticonvulsant, antiarthritic, osteopathic, ophthalmological and	
CC	cycostatic activities. This polypeptide sequence is a human heart	
CC	mitochondrial protein of the invention.	
XX		
SQ	Sequence 2263 AA;	

```

Alignment Scores:
Pred. No.:      2,5e-13      Length:      2263
Score:          301.50      Matches:      154
Percent Similarity: 37.5%      Conservative: 79
Best Local Similarity: 24.8%      Mismatches:  234
Query Match:      10.6%      Indels:      156
DB:              7          Gaps:        28

US-10-665-990A-13 (1-1561) x ADJ07425 (1-2263)

Qy      43  AACACGGAGCCTCATTTCCCTTTTATGGCTCTCTCTGTTCAATGTTCTTCATGTTGCC 102
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      358 SerThrGlyProGluProProAlaProThrProLeuLeu----- 370
Qy      103 CCCACTGGAGAACGGACGGAGAAACCGCGTCATTTCGAATACTTCCAAACCTGTCTCCTCGGA 162
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      371 ---AlaGluArgHisGlyGlySerPro-----GlnProLeuAlaThrProLeuSer 387
Qy      163 CAACATCTCTGCAAAATCG-----GCACACCCCTCATTAACAACGGGCTATCGACAT 213
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      388 GlnGluProValAlaGlnProSerSerGluAlaSerProThrArgAspArgSerProProLys 407
Qy      214 CTACTCTGCTCAGCACGCC-----CCACGAAGCCCTTTGCCGC 249
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      408 SerProGluLysLeuProGlnSerSerSerSerGluSerSerProSerPro---Gln 426
Qy      250 CCGCGCGCCCTTATCGAATCTGCGGAACACAGCCTCGATTGTCATTACTACATTGGGG 309
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      427 ProThrLysValSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446
Qy      310 CAACGACATTTCCGGCAGGCTGTTGTTCAAACCT----- 342
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```
Db 447 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
QY 343 -----CATGTACTTCCCGAGAACCGCGCTGCGGTACG 378
Db 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
QY 379 CCGTGTGTGGACGACACACACCGCGGGTGTGACGATCTCT- 423
Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
QY 424 GTCCTCCCTCGACACGCAATCCCAATATCGAAGTGGCGCTGTTCACACCCCTTCGTCCTACG 483
Db 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerAlaGln- 520
QY 484 CAAATGGCGCGACTCGGTACTACGACGACTTCCCGCCCTCAACCGCCGATGCACAA 543
Db 521 ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro- 537
QY 544 CAAATCCTTTACCGCGGACACCGCGCCACCATATCTCGGCGACGCAATATCGGCGACA 603
Db 538 -----GlnArgProGlyTrpSerArgSerArgSerArgSerGlnArgArgGlyArg 553
QY 604 ATACTTCAAAGTCGTGAGCACCGCTTTTCGCCGACCTGGACATCTCGCCACCGCGAG 663
Db 554 SerArgSerAlaArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
QY 664 CTTCTCGCGGAAGTATCGCACGACTTCGACCGTACTGGCAACCCATTCGCCCCACAA 723
Db 574 ArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAlaArgArgArg 593
QY 724 CGCCACGGCATCATCCGACGGCAACAT-----CGCAAGG 762
Db 594 SerArgSerArgThrProThrArgArgArgSerArgSerArgThrProAlaArgArgGly 613
QY 763 TCTTCAAGCACTCGGATA---CAACGACGAAATCCAGACACGCGCTCTCGCTACCG 819
Db 614 ArgSerArgSerArgThrProAlaArgArgArgSerArgSerArgThrPro---ValArg 632
QY 820 CGAAACCGTTCGAAGTCGCCCTTACCAAAAATACAGACGGACGCATCGACTGGCA 879
Db 633 ArgArgSerArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro 652
QY 880 GAGCGTCCAAACCCGCTGATCAGCACACCCCTGCMAA-----AGGACTGACCGGCA 933
Db 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
QY 934 CGCCGCAAAACCGCGATTCGCGGAGGCTGCAAGA----- 969
Db 673 ThrProAlaArgArgSerGlyArgSerArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
QY 970 CGCGTCAAAAGCCCGCAAAAAGCGTCTATCGTTTACCCCTATTTCTGTCCTCCACAA 1029
Db 693 ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg 712
QY 1030 ATCCCGCACAGA-----CGCACTGGCAAACTGTGTCAGGACGCGATAGAGTTAC 1080
Db 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerSerGluArgLys--- 731
QY 1081 CGTCTGACCAACTCGTACAGCGACCGACTTCGCGC----- 1119
Db 732 -----AsnLysSerArgThrSerGlnArgArgSerArgSerArgSerProGluMet 749
QY 1120 -----CGTCCATTCCGGTACGTCFAAATACCGAAACCGCGTGT 1158
Db 750 LysLysSerArgIleSerSerArgSerArgSerArgSerLeuSerSerProArgSerLysAla 769
QY 1159 CAAACCGCGCATCAACTCTAGA-----GCTGCAACCAACCATGCGCTCCCGCCAC 1212
Db 770 LysSerArgLeuSerLeuArgArgSerLeuSerGlySerSerProCys---ProLysGln 788
QY 1213 AAAAGCAAAAGCCCTGACCGGAGCTCCGTAACCA----- 1247
Db 789 LysSerGlnThrProProArgArgSerArgSerArgSerGlySerSerGlnThrLysAlaLysSer 808
```

```
QY 1248 -----CCCTGCATGCCAAAACCTTCAATTGTGGACGGCA 1280
Db 809 ArgThrProProArgArgSerArgSerSerSerProProLysGlnLysSerLys 828
QY 1281 AACGCATCTTCACTCGGCTCATTTCAACCTCGACCCCGTTCCG- 1322
Db 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
QY 1323 -----CACGGCTCAATACCGAAATGGCGCTGTCATCGAAGCCCAAAATTCGCAG 1373
Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluGlnSerValThrProGln 868
QY 1374 AACAGATGAGGAGCACCTCCCGCATACACACCGAATACGCTACCGCGTTACCTTCG 1433
Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
QY 1434 ACAACACACACCGCTGCAATGGC-----ACGATCCCGCA 1469
Db 888 ArgHis-----SerCysSerGlySerSerProProArgValLysSerSerThrProPro 905
QY 1470 CCCGAAATAACCTACCCGA---ACGAACCCGAAGCCA----- 1502
Db 906 ArgGlnSerProSerArgSerSerSerProGlnProLysValLysAlaIleSerPro 925
QY 1503 -----AAGTTGGAAACGATCGCGCGCAAAATCTATCCCTGCTGCCCATCG 1550
Db 926 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 939
QY 1551 AAGGTT 1556
Db 940 ArgVal 941

RESULT 40
ABO79612
ID ABO79612 standard; protein; 618 AA.
XX ABO79612;
AC ABO79612;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #11787.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX N-PSDB; ABD13183.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 28358; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
```


CC bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biotech technology. Sequences AB067826-AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX
SQ Sequence 618 AA;

Alignment Scores:

Pred. No.:	1,99e-13	Length:	618
Score:	301.00	Matches:	175
Percent Similarity:	34.6%	Conservative:	31
Best Local Similarity:	29.4%	Mismatches:	196
Query Match:	10.6%	Indels:	194
DB:	7	Gaps:	37

US-10-665-990A-13 (1-1561) x AB079612 (1-618)

QY	169	CCTGCAATTCGGACACACCCCTCATACAAACGGCTATCCGACATCTACCTGCTCGACA	228
DB	4	ProAspHisProArgHisProTrp--ArgArgGlnLeuAlaGluLeuArgIleArg	22
QY	229	-----CCCCACGAGCCCTTCCGCCCGCCGCCCTTATCGAATTCGCCGAACA	279
DB	23	ArgThrAlaProGlyGlyAspArgGlnGlnProArgArgThrProArgIleArgArg	42
QY	280	CAGCTCGATTGTCACTACTACATTGGCGCAACGACAT-----	318
DB	43	-----HisGlnGlyGlnSerHisArgProAlaHisGlnGlyArg	55
QY	319	-----TTCCGGCAGGCTGTCTTCAACCTCAT	345
DB	56	GlnArgGlnProGlyAlaProAlaAlaMetGlyLeuArgArgAla-----TrpProAla	73
QY	346	GTACCTTCGGCAGAACCGGGTGGCGGTAGCGCTGTGTGGACGACAAACACGCG	405
DB	74	AlaProCysGlyArgArgArgProAspThr-----GlnLeuArgLeu	88
QY	406	CGGGTTGGACCATCTCTGCTCGCCTCGACAGCATCCCAATATCGAAGTGGCCCTGTT	465
DB	89	ArgProGlnArgGlnSerGlyArgArgAsnGlnPro-----AlaProVal	103
QY	466	CAACCCCTT-----CGTCTTACGCAAAATGGCGGC	495
DB	104	ArgProGlnProGlyLeuArgArgProArgProAlaGlyArgProGluArgSerProArg	123
QY	496	ACTGGCTACTGACGACTTCCCGCCCTCAACCGCGCATGCACAAACAA-----	546
DB	124	ArgGlnAspProThrArgLeuArgArgProGlyGlnProHisArgGlyGlnGlyProAla	143
QY	547	-----ATCCTTTTACCGCGCAACCG-----	567
DB	144	ArgArgHisHisProLeuArgIleArgArgProArgProArgGlnProAspProThrGlyGlnPro	163
QY	568	-----CGCCACCATCTCGCGGACGCAATATATCGGCGACGAATACTTCAAGTCCGTGA	621
DB	164	GlyGlnArgHisHisLeuArgAlaArgArgArg-----GlnArgHisPro	180
QY	622	GGACACGTTTTCCGGACCTGGACATCTCGCACCGCGGACGCTGTGGCGAAGTATC	681
DB	181	ProHisArgArgSerArgCysGlyHis---ArgValSerLeuArgArgProGlnSerAla	199
QY	682	GCAGCACTTCGACCGCTACTGGCAAGCCATTTCGCCCCACAAACGCCACGCATCATCCG	741
DB	200	GlyArg---AlaProLeuAlaGluArgPro---GluProArgArgThrVal-----Pro	215

QY	742	CAGCGCAACATCGGCAAGGGTCTTCAAGCACTCGATACACGACGCAACATCCAGACA	801
DB	216	LeuArgProHisArg-----ArgArgGlnProGlyHis	226
QY	802	CGCGTCTCTGGCTACCGCAACCGTCAACAGTCGCCCTCTACTCAACAAATACACAG	861
DB	227	ArgProProGly-----ArgHisArgArg-----CysProArgGlnProGly	240
QY	862	GGGACGATCTACCTGGGAGACGGTCCAAACCGCTGTATCAGCGACACCCCTGCAAAAGG	921
DB	241	ValProLeuArgArgAlaArgGlnProGlyArgAlaGlyThrGlnHisPro-----	257
QY	922	ACTGACCGCGACCGCGCAACCGCGATTGCGGGAG-----GCTGCAAGCGCGCTCAA	978
DB	258	-----ProArgProAlaAspProAlaArgProArgAspLeuProLeuArgArgGlyGln	275
QY	979	ACAGCCCGAAAAAGCGTCTATCTGTTCACCCCTATTTCGTCTACAAATCCGGCAC	1038
DB	276	ProThrAlaGlyAspArgLeu-----ProLeuArgProArgHis	288
QY	1039	-----AGACGCACCTGGCAAA-----ACTGTCGAGGACGGCAT	1071
DB	289	ArgLeuProAlaGlnArgArgArgProGlyArgGlnArgAspProGlySerGlyArgGln	308
QY	1072	AGCGTTACCGTCTCT-----GACCAACTCGCTACA-----	1101
DB	309	GlyAlaGluHisProGlyTrpAlaAspArgLeuProAlaLeuArgProAlaAlaPro	328
QY	1102	-----GGCGACCGACGTTGCCGCTTCATTCGGCTACGTCAA-----ATACCGAAACCC	1152
DB	329	AspLeuGlyGlnArgHisSerGlnPro---ArgValArgProGlyLeuProThrAla	347
QY	1153	GCTCTCAAGCCCG-----CATCAAACTCTACAGCTGCAACCCAA-----	1194
DB	348	AlaAlaGluGlyArgProLeuAlaGluArgLeuProAlaArgCysGlnTrpGlnTrpPro	367
QY	1195	-----CCATCGCTCCCGCCACAAAGA	1218
DB	368	AlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro---	386
QY	1219	CAAGGGCTGACCGGACGCTCCGTAACCGCTCGATGCCAA---AACCTTTCATTGTGGA	1275
DB	387	-----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuArgLeuArgGly	404
QY	1276	CGCAAAACGCGAT-----CTTCATCGCTCATTCACCT	1308
DB	405	ArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---GlnPro	423
QY	1309	CGACCCCGTTCCGCGACGCTCAATAC-----CGAAATGGCGCTCGTCATCGAAAG	1359
DB	424	GlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgArgAlaSerGlyAsp	443
QY	1360	CCCCAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCCACCCGGAATACGCTTA	1419
DB	444	GlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArgArg	463
QY	1420	CGCGTTACCTCGACAAACACACCGCTCGCAATGCA-----	1458
DB	464	ProGlyProProAlaLysArgGlnProArgProAlaGlyArgArgIleProLeuGln	483
QY	1459	-----CGA	1461
DB	484	ArgProArgProAlaHisArgGlnAlaHisProArgGluHisHisLeuProLeuArg	503
QY	1462	TCCGCGCACCGCAAAACCTTACCGGAACCGAAGCCGAACTTTGGAAACCATCGC	1521
DB	504	SerArgArgProValAlaArgSerArgThrArgArgGlnArgProGluThrAlaSer	523
QY	1522	CGCAAAATCCTATC-----CCTGCTGCC-----CATCGAAGG	1554
DB	524	ProValLeuProValAlaGlyGlnProAlaAlaGlyHisHisArgArg	539

Search completed: May 2, 2006, 05:12:31
Job time : 390 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:52:55 ; Search time 49.2 Seconds
(without alignments)
2651.348 Million cell updates/sec

Title: US-10-665-990A-13

Perfect score: 2852

Sequence: 1 caaatacaggaatgcgct.....tgcccatcgaaaggtttatta 1561

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spo01/US10665990/runat_01052006_112001_9763/app_query.fasta_1
-DB=PublishedApplications_AA_Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10665990 @CGN_1_1_307 @runat_01052006_112001_9763 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA_Main:

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2689	94.3	525	5	US-10-665-990A-14
2	2681	94.0	525	5	US-10-665-990A-16
3	2677	93.9	525	5	US-10-665-990A-18
4	2630	92.2	525	4	US-10-066-551-4
5	2630	92.2	525	5	US-10-665-990A-4
6	2630	92.2	525	5	US-10-665-990A-20
7	632	22.2	502	4	US-10-335-977-9085
8	632	22.2	502	4	US-10-335-977-9086
9	590	20.7	428	3	US-09-881-752A-356
10	359	12.6	224	4	US-10-335-977-9084
11	336.5	11.8	5179	3	US-09-922-217-1068

12	336.5	11.8	5179	3	US-09-833-263-1068	Sequence 1068, Ap
13	336.5	11.8	5179	4	US-10-025-380-1068	Sequence 1068, Ap
14	336.5	11.8	5179	4	US-10-734-564-121	Sequence 121, App
15	323.5	11.3	1367	3	US-09-801-368-108	Sequence 108, App
16	310	10.9	563	4	US-10-437-963-198755	Sequence 198755, Sequence 46, Appl
17	303.5	10.6	2296	5	US-10-696-909A-46	Sequence 46, Appl
18	303.5	10.6	2752	5	US-10-696-909A-44	Sequence 44, Appl
19	301.5	10.6	2263	4	US-10-408-765A-2231	Sequence 2231, Ap
20	300	10.5	528	3	US-09-840-746-20	Sequence 20, Appl
21	298	10.4	497	4	US-10-437-963-125004	Sequence 125004, Sequence 5, Appli
22	297.5	10.4	19723	4	US-10-084-846A-5	Sequence 59, Appl
23	296	10.4	486	4	US-10-418-861B-59	Sequence 86, Appl
24	293.5	10.3	866	5	US-10-626-832-86	Sequence 174113, Sequence 176617, Sequence 3854, Ap
25	293	10.3	465	4	US-10-437-963-174113	Sequence 22, Appl
26	292.5	10.3	533	4	US-10-437-963-176617	Sequence 18, Appl
27	289.5	10.2	495	4	US-10-724-972A-3854	Sequence 18, Appl
28	286.5	10.0	493	3	US-09-966-521-22	Sequence 90, Appl
29	286.5	10.0	493	3	US-09-966-521-18	Sequence 84, Appl
30	286.5	10.0	493	4	US-10-429-094-18	Sequence 84, Appl
31	286.5	10.0	502	3	US-09-828-523A-90	Sequence 41094, A
32	286.5	10.0	502	3	US-09-966-521-84	Sequence 19493, Sequence 134726, Sequence 8, Appli
33	286.5	10.0	502	4	US-10-429-094-84	Sequence 9083, A
34	283	9.9	19695	4	US-10-084-846A-3	Sequence 36210, A
35	281	9.9	2284	6	US-11-097-143-41094	Sequence 168762, Sequence 50724, A
36	280.5	9.8	820	5	US-10-684-422-44	Sequence 203630, Sequence 177697, Sequence 44, Appl
37	279	9.8	427	4	US-10-437-963-199493	Sequence 19493, Sequence 134726, Sequence 8, Appli
38	277.5	9.7	668	4	US-10-437-963-134726	Sequence 9083, A
39	277.5	9.7	19608	4	US-10-084-846A-8	Sequence 168762, Sequence 50724, A
40	274.5	9.6	206	4	US-10-335-977-9083	Sequence 203630, Sequence 177697, Sequence 44, Appl
41	273.5	9.6	1795	6	US-11-097-143-36210	Sequence 36210, A
42	272	9.5	437	4	US-10-437-963-168762	Sequence 168762, Sequence 50724, A
43	268	9.4	1098	5	US-10-450-763-50724	Sequence 50724, A
44	266.5	9.3	555	4	US-10-437-963-203630	Sequence 203630, Sequence 177697, Sequence 44, Appl
45	265	9.3	621	4	US-10-437-963-177697	Sequence 177697, Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-10-665-990A-14
; Sequence 14, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neisseria gonorrhoeae
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae 1291
US-10-665-990A-14

Alignment Scores:
Pred. No.: 6.86e-172 Length: 525
Score: 2689.00 Matches: 520
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match:	94.3%	Indels:	0
DB:	5	Gaps:	0
US-10-665-990A-13 (1-1561) x US-10-665-990A-14 (1-525)			
QY	2	AAATACAGGCAATCGCTCTGAAACTATATCCCGATGAAACACAGCGACCTCATTTCC	61
DB	6	LysIleGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer	25
QY	62	CTTTTATGCTCTCTCTGTTTCATGTTCTTCTCATGTTGCCCGCCCACTGGAAGAACGACG	121
DB	26	LeuLeuCysLeuLeuLeuLeuCysSerCysSerSerTrpLeuProLeuGluGluArgThr	45
QY	122	GAAGCGCTCATTTCAATTAATCTCCAAACCTGCTCTCTGACAAACATCTTCAATTCGG	181
DB	46	GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg	65
QY	182	CACACCCCTCATAACACGGGTATCCGACATCTACCTGCTCGACGACCCCGACCAAGCC	241
DB	66	HisThrProHisAsnAsnGlyLeuSerAspIleTyrLeuLeuAspAspProHisGluAla	85
QY	242	CTTGCGCGCGCGCCCTTATCGAATCTGCGGAACACAGCCTCGATTGCAATACTAC	301
DB	86	LeuAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr	105
QY	302	ATTTGGCGCAACGATTTCCGCGAGGCTGTGTTCAACCTCATGTACCTTCCCGCAGAA	361
DB	106	IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaGlu	125
QY	362	CGGCGGTGCGGTACGCTGCTGTGGACGACACACACGCGGGTTGACCATCTC	421
DB	126	ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu	145
QY	422	CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCACACCCCTTCGCTTA	481
DB	146	LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu	165
QY	482	CCCAATGCGCGCACTCGGCTACTGACCGACTTCCCGCCCTCAACCGCGCGATGCAC	541
DB	166	ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMethis	185
QY	542	ACCAATCTTTACCGCGACACCGCGCCACCATCTCGCGGAGCGAATACCGGCAC	601
DB	186	AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp	205
QY	602	GAATACTTCAAAAGTCGGTGAGACACCGTTCCTGCGCGACCTCGACATCTCCGCCCGGC	661
DB	206	GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly	225
QY	662	ACGCTGTCGGCAAGTATCGCACCGACTTCGACCGCTACTGGCAAGCCATTCGCGCCAC	721
DB	226	SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlahis	245
QY	722	ACGCGCAGCGCATCATCCGACGCGCACATCTCGCAGGGTCTTCAGCACTTCGGATAC	781
DB	246	AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr	265
QY	782	AACGAGAAACATCCAGACACCGCTCTCTGCGCTACCGCGAAACCGTTCGAAACGTCGCC	841
DB	266	AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro	285
QY	842	CTCTACCAAAAAATACAGACGGGACGCATCGACTGCGAGAGCGTTCAAACCGCGCTGATC	901
DB	286	LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle	305
QY	902	ACGACACCCCTGCAAAAGGACTTCGACCGCGACCGCGCAACCGCGGAGG	961
DB	306	SerAspThrProAlaLysGlyLeuAspArgAspArgGlyProProIleAlaGlyArg	325
QY	962	CTGCAAGACGCGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC	1021
DB	326	LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal	345

QY	1022	CCTACAAAATCCGCGCACAGACGCACTGGCAAAAACCTGGTGCAGGACGGATAGACGTTACC	1081
DB	346	ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr	365
QY	1082	GTCCTGACCAACTCGCTACAGCGACCGACGTTGCCCGCGCTCCATTCCGGCTACGTCAAA	1141
DB	366	ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHiserGlyTyrValLys	385
QY	1142	TACCGAAAACCGCTGCTCAAAACCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATGCC	1201
DB	386	TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla	405
QY	1202	GTCCCGCGCACAAAGAGCAAAAGCGCTGACCCGCGCAGCTCCGTAACACAGCTGCATGCCAAA	1261
DB	406	ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys	425
QY	1262	ACCTTCATTGTGGACGCGCAAAACGCACTTTCATCGGCTCATTCACACTCGACCCCGTTCC	1321
DB	426	ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer	445
QY	1322	GCACGGCTCAATACCGAAATGGCGTCTGTCATCGAAAGCCCCCAAAATCGCAGAACAGATG	1381
DB	446	AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet	465
QY	1382	GAGCGACCCCTCGCGGATACACACCCCGAATACGCTACCGCTACCGGTTACCTCGACAAACAC	1441
DB	466	GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis	485
QY	1442	AACGCCCTGCAATGSCACGATCCGCCACCCGAAACCTTACCCGAAACGAAACCCGAAGCC	1501
DB	486	AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla	505
QY	1502	AAACTTTGGAAACGCAATCGCGCGCAAAAATCCTATCCCTGCTGCCCATCGAAGTTTATTA	1561
DB	506	LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu	525
RESULT 2			
US-10-665-990A-16			
; Sequence 16, Application US/10665990A			
; Publication No. US2004025322A1			
; GENERAL INFORMATION:			
; APPLICANT: Apicella, Michael A.			
; APPLICANT: Edwards, Jennifer L.			
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis			
; TITLE OF INVENTION: Infections			
; FILE REFERENCE: 17023-031001			
; CURRENT APPLICATION NUMBER: US/10/665,990A			
; CURRENT FILING DATE: 2003-09-19			
; PRIOR APPLICATION NUMBER: US 10/621,184			
; PRIOR FILING DATE: 2003-07-15			
; PRIOR APPLICATION NUMBER: US 10/066,551			
; PRIOR FILING DATE: 2002-01-31			
; PRIOR APPLICATION NUMBER: US 60/344,452			
; PRIOR FILING DATE: 2001-10-23			
; PRIOR APPLICATION NUMBER: US 60/310,356			
; PRIOR FILING DATE: 2001-08-06			
; PRIOR APPLICATION NUMBER: US 60/266,070			
; PRIOR FILING DATE: 2001-01-31			
; NUMBER OF SEQ ID NOS: 32			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 16			
; LENGTH: 525			
; TYPE: PRT			
; ORGANISM: Neisseria gonorrhoeae 1090			
US-10-665-990A-16			
Alignment Scores:			
Pred. No.:	2,368-171	Length:	525
Score:	2681.00	Matches:	518
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	94.0%	Indels:	0
DB:	5	Gaps:	0

US-10-665-990A-13 (1-1561) x US-10-665-990A-16 (1-525)

```
QY 2 AAAATACAGGGAATCGCTGTAAGACTATATCCCGATGAAACACGAGCCTCATTTCC 61
Db 6 LysileGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer 25
QY 62 CTTTATGCTCTCTCTCTGTTTCATGTTCTTCTATGTTGCCCCCACTGGAAGACGGAGC 121
Db 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProLeuGluGluArgThr 45
QY 122 GAAAGCCGTATTCATTTCAATCTTCAACCTGCTCTCTGGACAAACATCCTGCAATCCGG 181
Db 46 GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTCATACAAACGGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAGCC 241
Db 66 HisThrProHisAsnAsnGlyLeuSerAspIleTyrLeuLeuAspAspProHisGluAla 85
QY 242 CTGCGCGCCGCGCGCTTATCGAATCTGCGGAACACAGCCTCGATTTCGAATACTAC 301
Db 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCCGACAGCATTTCCGGCAGGCTGTGTTCAACCTCATGTATCTTCCCGCAGAA 361
Db 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaGlu 125
QY 362 CGCGCGTGGCGGTACGCTGCTGTGGACGACACACACGCGCGGGTGGAGCATCTC 421
Db 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTGTTCACACCCCTTCGCTCA 481
Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAAAATGGCGGCACTCGGCTACTGACCGACTTCCCGCCCTCAACCGCGCATGCAC 541
Db 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMetHis 185
QY 542 AACAAATCTTTACCGCGCAACCGCGCACCATCTACCGCGCAGCCCAATATCGGCGAC 601
Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAAGTCGGTGAGCACACCGTTTTCGCGCACCTGGACATCTCCGCCACCGC 661
Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTGTGGCGGAAGTATCGCAGCTTCGACCGCTACTGGGCAAGCCATTTCGCCCGAC 721
Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCACGCATCATCCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATAC 781
Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGACGAAACATCAGACACCGCTCTGCGCTACCGCGAAACCGTCGAACAGCTCGCCC 841
Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCCTACCAAAAAATACAGAGGGAGCATCGACTGGCAGAGCGTCCAAAACCGCGCTGATC 901
Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATTCGCGGAGG 961
Db 306 SerAspSerProAlaLysGlyLeuAspArgAspArgLysProIleAlaGlyArg 325
QY 962 CTGCAAGACGCGCTCAACACAGCCGAAAAAGCGCTCTATCTGTTTCACCCCTATTTCGTC 1021
Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CCTTCAAAATCCGCGCACAGACGCACTCGGCAAAATCTGTGAGGACGCGATAGCGTTACC 1081
```

```
Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCTGTGACCAACTCGCTACAGCGACCGACCTTCCGCGCTGCATTTCCGGCTACGTCAA 1141
Db 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAAAACCGCTGCTCAAAGCGGCATCAAACTCTACGAGCTGCAACCCCAACCATGCC 1201
Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
QY 1202 GTCCCCCGCCAAAAAGAGCTGACCGGAGCTCCGTAACAGCCTCGATGCCCAAA 1261
Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTGGACGCGAAACGCGATCTTCATTCGGCTCATTAACCTCGACCCCGTTCC 1321
Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCACGGCTCAATACCGAAATGGGCTGCTCATCGAAAGCCCAAAATCGCAGAACAGATG 1381
Db 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGATACACACCGGAATACGCTACCGGTTACCTCGACAAACAC 1441
Db 466 GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis 485
QY 1442 AACCGCTGCAATGGCAGCATCCCGCCACCGCAAAAACCTACCGAAGAACCCGAGGCC 1501
Db 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTTGGAACGCGCATCGCGCAAAATCTATCCCTGCTGCGCCATCGAAGGTTTATTA 1561
Db 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu 525
```

RESULT 3

```
US-10-665-990A-18
; Sequence 18, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Ne
; TITLE OF INVENTION: infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae ms11
US-10-665-990A-18
```

```
Alignment Scores:
Pred. No.: 4,37e-171 Length: 525
Score: 2677.00 Matches: 517
Percent Similarity: 99.6% Conservative: 1
Best Local Similarity: 99.4% Mismatches: 2
Query Match: 93.9% Indels: 0
DB: 5 Gaps: 0
```

US-10-665-990A-13 (1-1561) x US-10-665-990A-18 (1-525)


```
Db      26  LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProProLeuGluGluArgThr 45
      122  GAAAGCCGCTCAATTCATCTTCCAAACCTGTCCTCCCTGGACAAACATCTCGAAATCCGG 181
      46  GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
      182  CACACCCCTCATAAACAGGGCTATCCGACATCTACCTGCTCGACAGACCCCCACAGAGCC 241
      66  HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
      242  CTTGCGCGCGCGCGCTTATCGAATCTGCGACACAGCCTCGATTGCAATACTAC 301
      86  PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
      302  ATTGGCGCAACACGATTTCCGGCAGGCTGTGTTCAACCTCATGTACCTCCCGCAGAA 361
      106  IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
      362  CGCGGGTGGCGGTACGCGCTGTGTTGGACGCAACACACGCGCGGGTGGACGATCTC 421
      126  ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
      422  CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTCAACCCCTTCGTCCTA 481
      146  LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
      482  CGCAATGGCGCGCACTCGGCTACCTGACCGACTTCCCGCGCTCAACCGCGCGATGCAC 541
      166  ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgMetHis 185
      542  AACAAATCTTTACGCGCGACAAACCGCGCCACCATCTCGGCGGAGGCAATATCGCGAC 601
      186  AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
      602  GAATACTTCAAGTCGGTGAGACACCGTTTTCGCGCGCTCGACATCTCGCCACCGCGC 661
      206  GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
      662  AGCGTCGTGGGGAAGTATCGCACGACTTCGACCGCTACTGGCGAAGCCATTCGCGCCAC 721
      226  SerValValGlyValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
      722  AACGCCACGCGCATCATCGCGAGCGGCAACATCGGCAAGGGTCTTCAAGCATCTCGGATC 781
      246  AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
      782  AACGAGAAACATCCAGACACGCGCTCTCGCGCTACCGGAAACCGTGGAAACAGTCGCC 841
      266  AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
      842  CTCCTACAAAAATACAGACGGGACGCATCGACTGCGAGAGCGTCCAAACCCGCTGATC 901
      286  LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
      902  ACGCACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAAAACCGCGGATTCGCGGAGG 961
      306  SerAspAspProAlaLysGlyLeuAspArgAspArgArgLysProProIleAlaGlyArg 325
      962  CTGCAAGACGCGCTCAAAACAGCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC 1021
      326  LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
      1022  CCTACAAAATCCGGCACAGACGACTGGCAAAACTGGTCCAGACGGCATAGACGTTACC 1081
      346  ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
      1082  GTCTTGACCAATCGCTACAGGGACCGAGTTCGCGCGCTCCATTCGGCTACGTCAAA 1141
      366  ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
      1142  TACCGAAAAACCGCTGTCTAAAGCCGGCATCAAACTCTACGAGCTGCAACCCCAACATGCC 1201
```

```
Db      386  TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
      1202  GTCCCCGCCACAAAGACAAAGCGCTGACCGGACAGTCCGTAACAGAGCTGCATGCCAAA 1261
      406  ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
      1262  ACCTTCATTGTCGAGCGCAAAAGCATCTTCATCGGCTCATCAACCTCGACCCCGCTTCC 1321
      426  ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
      1322  GCACGGCTCAATACCGAAATGGCGTCGTCTATCGAAAGCCCCAAAATCGCAGAACAGATG 1381
      446  AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
      1382  GAGCGCACCTCGCGCATACCAACCCGGAATACCGCTACCGCTACCGGTTTACCTCGACAAA 1441
      466  GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
      1442  AACCGCTGCAATGGCAGCATCCCGCCACCGCAAAACCTACCCGAAACGACCCGAGCC 1501
      486  AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
      1502  AAACCTTTGGAAACGCGATCCCGCAAAAATCCTATCCCTGCTGCCCATCGAAGGTTTATTA 1561
      506  LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525

RESULT 6
US-10-665-990A-20
; Sequence 20, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; TITLE OF INVENTION: Infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria meningitidis B
US-10-665-990A-20

Alignment Scores:
Pred. No.: 6,21e-168 Length: 525
Score: 2630.00 Matches: 509
Percent Similarity: 98.7% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 7
Query Match: 92.2% Indels: 0
DB: 5 Gaps: 0

US-10-665-990A-13 (1-1561) x US-10-665-990A-20 (1-525)
QY 2 AAAATACAGGCAATCCGCTCTGAAACTATATCCCCGATGAAAAACACGACGCGCTCATTTCC 61
DB 6 LysThrGlnAlaMetProSerGluThrIleSerLeuMetLysThrArgSerLeuLeuSer 25
QY 62 CTTTATGCTCTCTCTCTGTTCATGTTCATGTTCTTCATGTTGCCCCCTCGAAGACGAGC 121
DB 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProProLeuGluGluArgThr 45
```


QY 122 GAAAGCGCTCAATTCCTCAAACTGCTCTCCCTGGACAAACATCTGCGAAATCCGG 181
DB 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleuGlnIleArg 65
QY 182 CACACCCCTCAATAACAGGGGTATCCGACATCTACCTGCTCGACGACCCCAAGGCC 241
DB 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
QY 242 CTTCGCGCGCGCGCTTATCGAATCTGCGACACAGCTCGATTTCGATCTAC 301
DB 86 PheAlaAlaArgAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyr 105
QY 302 ATTTGCGCAACGACATTTCCGCGAGGTGCTGTTCAACTCTACCTCTGCGCGCAAA 361
DB 106 IleTyrArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaGlu 125
QY 362 CGCGCGTGGGTACGCTGCTGTGGACGACAAACACACGCGGGTGGACATCTC 421
DB 126 ArgGlyValArgValArgLeuLeuAspAspAsnAsnThrArgGlyLeuAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTGTTCAAACCCCTTCGCTCA 481
DB 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATGCGCGCACTCGGCTACCTGACCGCACTTCCCGCGCTCAACCGCGCATGCAC 541
DB 166 ArgLysTyrArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgMethHis 185
QY 542 AACAAATCTTTACCGCGACAAACCGCGCCACCTATCTCGGCGACGAATATCGGAC 601
DB 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAGTCGGTGAGACACCGCTTTTCGCGACCTGGACATCTCGCCACCGGC 661
DB 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTCTGCGCGAAGTATCGACGACTTCGACCGCTACTCGGCAAGCATTCGCGCCAC 721
DB 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCGACGCGATCATCCGCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATAC 781
DB 246 AsnAlaThrArgIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGCGAAACATCCAGACGCGCTCTCGCTACCGCAACCGTGAACACGTCGCGCC 841
DB 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCTACCAAAAAATACAGACGGGCGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATC 901
DB 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCTCTGAAAGAGTCTGACCGCGACCGCGCAACCGCGATTCGCGGAGG 961
DB 306 SerAspAspProAlaLysGlyLeuAspArgAspArgLysProIleAlaGlyArg 325
QY 962 CTGCAAGCGCTCAACACGCGCGCAAAAGCGCTCTATCTGGTTTACCCCTATTTCGTC 1021
DB 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CCTCAAAATCCGCGACAGCGCACTGGCAAACTGGTTCGAGGACGGCATAGAGCTTACC 1081
DB 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCCTGACCAACTCGCTACAGGACCGACGCTTCGCGCGCTCCATTCGGCTACGTCAAA 1141
DB 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaHisSerGlyTyrValLys 385
QY 1142 TACGCAAAACCGCTCAAGCGCGGATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
DB 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405

QY 1202 GTCCCGCGCACAAAGAGCGCTGACCGGAGCTCCGTAAACAGCGCTCATGCCCAA 1261
DB 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTGGACGCAACACGCATCTTCATCGGCTCAATCAACCTCGACCCCGTTCC 1321
DB 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCAGGCTCAATACGGAATGGCGTCTGCTCATCGAAAGCCCAAAATCGCAGAACAGATG 1381
DB 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGCATACACACCGCAATACCGCTACCGCTTACCTCGACAAACAC 1441
DB 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
QY 1442 AACCGCTGCAATGGCAGATCCCGCCACCGCAAAACCTACCCGAAACGAAACCGGAGCC 1501
DB 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTTGGAAACGATCGCGCGCAAAATCTCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
DB 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525

RESULT 7
US-10-335-977-9085
; Sequence 9085, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9085:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...502
; SEQUENCE DESCRIPTION: SEQ ID NO: 9085:
US-10-335-977-9085

Alignment Scores:
 Pred. No.: 6,25e-34 Length: 502
 Score: 632.00 Matches: 161
 Percent Similarity: 51.7% Conservative: 93
 Best Local Similarity: 32.8% Mismatches: 181
 Query Match: 22.2% Indels: 56
 DB: 4 Gaps: 17

US-10-665-990A-13 (1-1561) x US-10-335-977-9085 (1-502)

```

QY 176 ATCCGGCACACCCTCATAAACAGGGCTATCCGACATCTAC----- 217
Db 30 lIeSerTyAspProTyThrThrThrleGlySerLeuTyAlaLysAsnLeuLysGlu 49
QY 218 -----CTGCTCGACGACCCCCACGAGCCCTTCGCCCGCGC 253
Db 50 AsnProLysHisSerAlaAlaLeuLeuGluAspGlyPheAspAlaLeuHisArg 69
QY 254 GCGGCCCTTATCGAATCTCCGAACACACAGCCTCGATTGCAATACTACATTTGGCGCAAC 313
Db 70 ValGlyLeuIleArgMetSerGlnLysSerIleAspMetGlnThrTyrlleTyrlLysAsn 89
QY 314 GACATTTCCGGCAGGCTGCTTCAACCTCATGCTACCTTGGCCGAGAACGCGCGCTGCGC 373
Db 90 AspLeuSerSerGlnValIleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLys 109
QY 374 GTACGCTCTGCTTGGACGACAAACACGCGCGGGTTGGAC-----GATCTC 421
Db 110 ValArgIleLeuLeuAspAspAsn-----GlyLeuAspSerAspPheSerAspIle 126
QY 422 CTGCTCGCCCTCGACAGCCTCCCAATATCGAAGTGGCGCTTTCAACCCCTTCGTCCTA 481
Db 127 MetLeu---LeuAsnPheHisLysAsnIleGluValLysIlePheAsnProTyrlle 145
QY 482 CCACAAATGGCGCCTCGGCTAC-----CTGACCGACTTCCCGCGCTCAACGCG 532
Db 146 ArgAsn---LysGlyLeuArgTyrlPheGluMetLeuAlaAspTyrlGluArgIleLys 164
QY 533 CGCATGCACAAATCTTTTACC CGCGACAAACCGCGCCACCATACTCGGCGAGCGCAAT 592
Db 165 ArgMetHisAsnLysLeuPheIleValAspAsnPheAlaValIleIleGlyGlyArgAsn 184
QY 593 ATCGGGCAGCAATACTTCAAAGTCGGTGAGGACACCGTTTTCGCGCAGCTGCACATCTC 652
Db 185 IleGlyAspAsnTyrlPheAspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeu 204
QY 653 GCCACCGCGAGGCTCGTCGGCGAAGTATCCGACGACTTCGACGCTTACGCGCAAGCCAT 712
Db 205 PhePheGlyGlyValAlaSerLysAlaLysGluSerPheGluAsnTyrlTrpArgPheHis 224
QY 713 TCCGCCCAACACGCCACGCGCATCTCCGAGC-----GGC 748
Db 225 ArgSerIleProValSer---LeuLeuArgThrHisLysArgLeuLysAsnAsnValLys 243
QY 749 ACATCGGCAAGGCTTCAAGCCTCGATACACGACGAAACATCCAGACGCGCTC 808
Db 244 GluIleAlaLysLeuHisGluLysIleProIleSerAlaGluAspAlaAsnGluPheGlu 263
QY 809 CTGCGCTACCGGCAACCGCTCAACAGTCCGCCCTCTACCAA---AAATACAGCGGA 865
Db 264 LysLysValAsnAspPheIleGluArgPheGlnLysTyrlGlnTyrlProIleTyrlGly 283
QY 866 CGCATCGATGCGAGAGCTCCAAACCCGCTGATCAGGACACCCCTGCAAAAGACTC 925
Db 284 -----AsnAlaIlePheLeuAlaAspLeuProAlaLys---Ile 295
QY 926 GACCGCGACCGCGCAACCGCGATTGCGGGAGGCTGCAAGCGCGCTCAACAGCCCC 985
Db 296 AspThrProLeuTyrlSer---ProIleLysIleAlaPheGluLysAlaLeuLysAsnAla 314
QY 986 GAAAAAGCGCTTATCTGTTTCAACCTATTTTCGCTCCCTTACAAATCCGCGCAGACGCA 1045
  ::::: ||||| ::::: ||| ||||| ::::: |||

```

```

Db 315 LysAspSerValPheIleAlaSerSerTyrlPheIleProGlyLysLysIleMetLysIle 334
QY 1046 CTGCAAAACCTGGTGCAGGACGGCATAGACGTTACCTGCTCCTGACCAACTCGCTACAGGG 1105
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 335 PheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeuThrAsnSerLeuSer 354
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1106 ACCGAGGTTCGCGCGCTCATTCGGCTACGTCAAAATACCGAAAAACCCCTGCTCAAAGCC 1165
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 355 ThrAspAlaIleValValTyrlGlyAlaTrpGluArgTyrlArgAsnLysLeuValArgMet 374
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1166 GGCATCAAACTCTAGAGCTGCAACCCCAACCATGCGTCCCGCCGACAAAAAGACAAAGGC 1225
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 375 GlyAlaAsnValTyrlGluIleArgAsnAspPhePheAsnArgGlnIleLysGlyArg--- 393
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1226 CTGACCGGCGAGCTCGGTAAACAGCCTGATCGCAACCAACCTTCATTTGTGTGCGCGCAACGC 1285
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 394 -----PheSerThrLysHisSerLeuHisGlyLysThrIleValPheAspAspAlaLeu 411
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1286 ATCTTCATCGGCTCATTCCAACTCGACCCCGTTCGCGCAGCGCTCAATACCGAAATGGGC 1345
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 412 ThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyrlleAsnThrGluSerAla 431
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1346 GTCTGTCATCGAAAGCCCAAAATCGCAACAGATGGAGCGCGCCCTCGCGCATCCACA 1405
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 432 ValLeuPheAspAsnProSerPheAlaLysArgValArgLeuSerLeuLysAspHisAla 451
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1406 CCGAATACGCTACCGGTTACCTCGACAAAACAAACCGCTGCAATGGCAGCATCCC 1465
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 452 ---GlnGlnSerTrpHisLeuValLeuTyrlArgHis---ArgValIleTrpGlu----- 467
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1466 GCCACCCGAAAAACCTAC-----CCGAAACGAAACCGAAGCCCAACTTTTGAAA 1513
  ||||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
Db 468 AlaThrGluGluGlyIleLeuIleHisGluLysAsnSerProAspThrSerPhePheLeu 487
  ::||| ::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
QY 1514 CGCATCGCGCAAAATCTATCCCTGCTGCC 1546
  ::||| ::::: ||||| ::::: ||||| ::::: |||||
Db 488 ArgLeuIleLysGluTrpSerLysValLeuPro 498
  ::||| ::::: ||||| ::::: ||||| ::::: |||||

```

RESULT 8
 US-10-335-977-9086
 ; Sequence 9086, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ;
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: UNIX
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/335,977
 ; FILING DATE: 30-Dec-2002
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/993,002
 ; FILING DATE: 17-DEC-1997
 ;
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragouras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: GTN-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 742-4214

; INFORMATION FOR SEQ ID NO: 9086:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 502 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Helicobacter pylori
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...502
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 9086:
 US-10-335-977-9086

Alignment Scores:

Pred. No.: 6, 25e-34 Length: 502
 Score: 632.00 Matches: 161
 Percent Similarity: 51.7% Conservative: 93
 Best Local Similarity: 32.8% Mismatches: 181
 Query Match: 22.2% Indels: 56
 DB: 4 Gaps: 17

US-10-665-990A-13 (1-1561) x US-10-335-977-9086 (1-502)

```

Qy 176 ATCCGGCACCCCTCTATACAAACGGGCTATCCGACATCTAC----- 217
Db 30 lIeSerTyAspProTyThrThrThrleGlySerLeuTyAlaLysAsnLeuLysGlu 49

Qy 218 -----CTGCTCAGACNCCCCAGAACGCCCTTTCGCCCGCGC 253
Db 50 AsnProLysHisSerAlaAlaileLeuLeuGluAspGlyPheAspAlaLeuHisArg 69

Qy 254 GCCGCCCTTATCGAATCGCCGACACAGCTCGATTTGCAATACTACTATTGGCGCAAC 313
Db 70 ValGlyLeuileArgMetSerGlnLysSerileAspMetGlnThrTyrlleTyLysAsn 89

Qy 314 GACATTTCCGGCAGGCTGCTTCAACCTCATCTGTCGCGAGAACGGGGTGCGC 373
Db 90 AspLeuSerSerGlnValleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValys 109

Qy 374 GTACGCTGCTGTTGGACGACAAACACACGGCGGGTGGAC-----GATCTC 421
Db 110 ValArgileLeuLeuAspAsn-----GlyLeuAspSerAspPheSerAspille 126

Qy 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTTCAACCCCTTCGTCTA 481
Db 127 MetLeu---LeuAsnPheHisLysAsnilleGluValLysilePheAsnProTyTyrlle 145

Qy 482 CGCAATATGGCGCACTCGGCTAC-----CTGACCGCACTTCCCGCGCTCAACCGC 532
Db 146 ArgAsn---LysGlyLeuArgTyThrPheGluMetLeuAlaAspTyThrGluArgileLysLys 164

Qy 533 CGATGACAAACAATCTTATCCGCGGACAAACCGCGCACCATCTCGCGGAGCGCAAT 592
Db 165 ArgMetHisAsnLysLeuPheileValAspAsnPheAlaValillelleGlyArgAsn 184

Qy 593 ATCCGCGACGAATCTCAAGTCGGTGAGACACCGTTTTCGCGGACCTGGACATCTC 652
Db 185 lIeGlyAspAsnTyThrPheAspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeu 204

Qy 653 GCCACCGCGACGCTCGTGGCGGAAGTATCGCACGACTTCGACGACTTCCGACGAGCAT 712
Db 205 PhePheGlyGlyValAlaSerLysLysAlaLysGluSerPheGluAsnTyThrArgPheHis 224

Qy 713 TCCGCCCAACGCCGCGCATCATCCGACG-----GGC 748
Db 225 ArgSerileProValSer---LeuLeuArgThrHisLysArgLeuLysAsnAsnValys 243

Qy 749 AACATCGCGAGGCTCTCAAGCACTCGGATACACGACGAAACATCCAGACGCGCTC 808
Db 244 GluileAlaLysLeuHisGluLysileProilleSerAlaGluAspAlaAsnGluPheGlu 263
  
```

```

Qy 809 CTGGCTACCGCGAAACCGTCAACAGTCGCGCCCTCTACCAA---AAAATACAGACGGGA 865
Db 264 LysLysValAsnAspPheilleGluArgPheGlnLysTyThrGlnTyThrProilleTyThrGly 283

Qy 866 CGCATCGACTGGCAGAGCTCCAAACCGCTGTATCAGCAGACCCCTGCAAAAGGATCTC 925
Db 284 -----AsnAlailePheLeuAlaAspLeuProAlaLys---ile 295

Qy 926 GACCGCGACCGCGCAACCGCATTCGCGGGAGGTGCAAGACGCGCTCAACAGACGCC 985
Db 296 AspThrProLeuTyThrSer---ProilleLysileAlaPheGluLysAlaLeuLysAsnAla 314

Qy 986 GAAAAAGCGTCTATCTGGTTTTCACCCCTATTTCCCTCAAAATCCGACACGACGCA 1045
Db 315 LysAspSerValPheilleAlaSerSerTyThrPheilleProGlyLysLysileMetLysile 334

Qy 1046 CTGGCAAAACTGTGTGAGGAGCGATAGACGTTACCGTCTCTGACCAACTCTCAGCGG 1105
Db 335 PheLysAsnGlnilleSerLysGlyileGluLeuAsnilleLeuThrAsnSerLeuSerSer 354

Qy 1106 ACCGACGTTGCCGCGCTCCATTCGCGCTACGTCAATACCGAAACCGCTCTCAAGGCC 1165
Db 355 ThrAspAlaileValValTyThrGlyAlaTrpGluArgTyThrArgAsnLysLeuValArgMet 374

Qy 1166 GGCAATCAAACTCTACGAGCTGCAACCAACCATCGCTCCCGCCGACAAAGACAAAGGC 1225
Db 375 GlyAlaAsnValTyThrGluileArgAsnAspPhePheAsnArgGlnilleLysGlyArg--- 393

Qy 1226 CTGACCGGCGAGTCCGTAAACAGCTGCATCGCCGCTCAACAACTTCATTTGGACGCAACGC 1285
Db 394 -----PheSerThrLysHisSerLeuHisGlyLysThrileValPheAspAlaLeu 411

Qy 1286 ATCTTCATCGCTCATTCACCTCGACCGCTCGCCGCTCGCACGGCTCAATACCGAAATGGC 1345
Db 412 ThrLeuLeuGlySerPheAsnilleAspProArgSerAlaTyThrileAsnThrGluSerAla 431

Qy 1346 GTCTCATCAAAAGCCCCAAATCGCAGAACAGATGAGGCGACCCCTCGCGATACCAACA 1405
Db 432 ValLeuPheAsnProSerPheAlaLysArgValArgLeuSerLeuLysAspHisAla 451

Qy 1406 CCGCAATACGCTTACCGCGTTACCTCTCGACAAACACACCGCTCGCAATTCGACGATCCC 1465
Db 452 ---GlnGlnSerTrpHisLeuValLeuTyThrArgHis---ArgValilleTrpGlu----- 467

Qy 1466 GCCACCGCGAAACCTAC-----CCGAACGAAACCGGAAGCAAACTTTGGAAA 1513
Db 468 AlaThrGluGluGlyLysileLeuLysGluLysAsnSerProAspThrSerPhePheLeu 487

Qy 1514 CGCATCGCGCAAAATCTTATCCCTGCTGCC 1546
Db 488 ArgLeuileLysGluTrpSerLysValLeuPro 498
  
```

RESULT 9

US-09-881-752A-356
 ; Sequence 356, Application US/09881752A
 ; Patent No. US20020115078A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kleanthous, Harold
 ; APPLICANT: Al-Garawi, Amal
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Tomb, Jean-Francois
 ; APPLICANT: Oomen, Raymond P.
 ; TITLE OF INVENTION: Identification of Polynucleotides
 ; TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in the
 ; FILE REFERENCE: 06132/041002
 ; CURRENT APPLICATION NUMBER: US/09/881,752A
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 08/833,457
 ; PRIOR FILING DATE: 1997-04-01
 ; NUMBER OF SEQ ID NOS: 370
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 356

```
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-356

Alignment Scores:
Pred. No.:      3,998-31      Length:      428
Score:          590.00      Matches:      145
Percent Similarity: 50.1%      Conservative: 80
Best Local Similarity: 32.3%      Mismatches: 146
Query Match:      20.7%      Indels:      78
DB:               3         Gaps:      14

US-10-665-990A-13 (1-1561) x US-09-881-752A-356 (1-428)
QY 272 GCGGACACAGCCTCGATTGCAATACATACATTTGGCGCAACAGACATTTCCGGCAGGCTG 331
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 SerGlnLysSerIleAspMetGlnThrTyrIleTyrLysAsnAspLeuSerSerGlnVal 21
QY 332 CTGTTCAACCTCATGTACCTTCCCGGAGAACCGCGGTGCGGTACGCTGCTGTGGAC 391
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 22 IleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLysValArgIleLeuLeuAsp 41
QY 392 GACAAACACACGCGGGGTGAC-----GATCTCTCTGCTCGCCTCGACAGC 439
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 42 AspAsn-----GlyLeuAspSerAspPheSerAspIleMetLeu----LeuAsnPhe 57
QY 440 CATCCCAATATCGAAGTGGCTGTTCACCCCTTCCTACCGCAAAATGGCGCGCACTC 499
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 58 HisLysAsnIleGluValLysIlePheAsnProTyrTyrIleArgAsn---LysGlyLeu 76
QY 500 GGCTAC-----CTGACCGACTTCCCGCCTCAACCGCGCATGACACAAATCC 550
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 77 ArgTyrPheGluMetLeuAlaAspTyrGluArgIleLysLysArgMetHisAsnLysLeu 96
QY 551 TTTACCGCGCAACCGCCACCATCTCGCGGAGCGCAATATCGGCGAGCAATACTTC 610
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 97 PheIleValAspAsnPheAlaValIleIleGlyArgAsnIleGlyAspAsnTyrPhe 116
QY 611 AAGTGTGAGACACACCGTTTCGCGGACCTGGACATCTCGCCACCGGCGAGCGTCGT 670
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 117 AspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeuPheGlyValAla 136
QY 671 GCGGAAGTATCCACAGCTTCGACCGCTACTGGGCAAGCCATTCGCCCCACACGCCACG 730
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 137 SerLysAlaLysGluSerPheGluArgTyrTrpArgPheHisArgSerIleProValSer 156
QY 731 CGCATCATCCGAGCGGCAACATCGGCAAGGCTCTCAAGCATCGGATACACACGACAA 790
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 157 ----LeuLeuArgThr----- 160
QY 791 ACATCCAGACACGCGCTCCTCGCTACCGCGAAACCGTGAACAGTCGCGCCCTTACCAA 850
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 161 -----HisLysArgLeuLys---AsnAsnAlaLysGluIleAlaLysLeuHisGlu 176
QY 851 AAATACACAGCGGAGCATCATCGGCGAGCGTCCAAACCGCGCTG----- 898
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 177 LysIleProIleSerAlaGluAspLysAsnGlnPheGluLysLysValAsnAspPheIle 196
QY 899 ----- 907
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 197 AspArgPheGlnLysTyrGlnTyrProIleTyrTyrGlyAsnAlaIlePheLeuAlaAsp 216
QY 908 ACCCTTGCAAAAGACTCGACCGCGACCGCGCAACCGCGATGTCGCGGAGGCTGCAA 967
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 217 SerProLysLys---IleAspThrProLeuTyrSer---ProIleLysIleAlaPheGlu 234
QY 968 GACGCGCTCAACAGCCCCGAAAGCGTCTATCTGGTTTCACCTATTTCTCCTCTACA 1027
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 235 LysAlaLeuLysAsnAlaLysAspSerValPheIleAlaSerSerTyrPheIleProGly 254
QY 1028 AAATCCGGCACAGACGCACTGGCAAAACCTGGTGCAGGACGGCATAGACGTTACCGTCTG 1087
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 255 LysLysMetMetLysIlePheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeu 274
QY 1088 ACCAACTCGCTACAGGCGACGCGATTGCCGCCGCTCCATTCGGGCTCAAGTCAAATACCGA 1147
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 275 ThrAsnSerLeuSerSerThrAspAlaIleValValTyrGlyAlaTyrGluArgTyrArg 294
QY 1148 AAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCAACCATGCGCTCCGC 1207
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 295 AsnGlnLeuValArgMetGlyAlaAsnValTyrGluIleArgAsnAspPhePheAsnArg 314
QY 1208 GCCACAAAGACAAAGCGCTGACCGGAGCTCCGTAACACGAGCTCATGCCAAACCTTC 1267
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 315 GlnIleLysGlyArg-----PheSerThrLysHisSerLeuHisGlyLysThrIle 331
QY 1268 ATTGTGGACGGCAACGCGATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGACGG 1327
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 332 ValPheAspAspAsnLeuThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyr 351
QY 1328 CTCAATACCGAAATGGGCGTCTCATCGAAAGCCCCCAAAATCGCAGACAGATGGAGCGC 1387
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 352 IleAsnThrGluSerAlaValLeuPheAspAsnProSerPheAlaLysArgVal----- 369
QY 1388 ACCCTCGCGGATACACACCCCGAATACGCTACCGCTTACCTCGACAAACACAAACGCG 1447
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 370 -----ArgLeuSerLeuLysAspHisAlaGln 378
QY 1448 CTGCAATGCGACGATCCGCGCACCGCAAAACCTTACCCGAAACGACCGAAGCCAACTT 1507
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 379 GlnSerTyrHisLeuValValTyrArg-----HisArgValIle 391
QY 1508 TGGAAACGCAATCGCGCGCAAAATCCTA 1534
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 392 TrpGluAlaValGluGluGlyLeu 400

RESULT 10
US-10-335-977-9084
; Sequence 9084, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9084:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
```



```
Db      1441 oileThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
Qy      581 GCGGACGCAATATCGCGCAGCAATACTTAAAGTCGGTGAGGACACCGGTTTTTCGCCGAC 640
Db      1457 -----ProIleSerThrThrThrProProProThrThrThrProSerProProTh 1474
Qy      641 CTGGACATCTCGCCACCGGAGCGTCGTCGGCGAAGTATCGCAGCATTCGACG----- 695
Db      1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
Qy      696 -----GCTACTGGCAAGCAATTCGCCGCACACGCCACCGCATCATCGGAGCGGC 748
Db      1494 oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh 1514
Qy      749 AACATCGGCAAGGGTCTTCAAGCACTCGGATACACGACGAAACATCCAGACAGCGCGTC 808
Db      1514 rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh 1529
Qy      809 CTGCGCTACCGGGAACCGTGAAC-----AGTCGCCCTCTACCAAAAAATACAGACG 862
Db      1529 rThrThrThrProProThrThrThrProSerProProThrThr----- 1544
Qy      863 GGACGCATCGTGCAGAGCGTCAAAACCGCCTGATCAGCGACACCCCTGCAAAAGGA 922
Db      1545 -----ThrProIleThrProPr 1550
Qy      923 CTGACCGCGACCGCGCAAAACCGCGGATTCGCGGAGGCTGCAAGACGCGCTCAAAACAG 982
Db      1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
Qy      983 CCGGAAAAAGGTTCTATCTGGTTTCACCTTATTCGTCCTTACAAATCCGGCAGAC 1042
Db      1563 rProProThrThrThrThrThrThrProProThrThrThrProSerProProThrTh 1583
Qy      1043 GCACGTGGCAAAACTGGTCAGCGCATAGAGTTACCGTCTGACCA-----AC 1093
Db      1593 r-----ThrThrProSerProProThrIleThrTh 1593
Qy      1094 TCGGTACAGGCGACGAGTTCGCGCGTTCATTCGCGCTACGTCAAAAT----- 1142
Db      1593 rThrThrProProThrThrThrProSerProProThrThrThrThrThrThrProPr 1613
Qy      1143 -----ACGAAACCGTGTCTCAAAAGCGGCATCAACTC 1177
Db      1613 oProThrThrThrProSerProProThrThrThrProIleThrProProThrSerThrTh 1633
Qy      1178 TACGAGCTGCAACCAACCATCGCTCCCGCCACAAAGACAAAGGCTGACCGGACG 1237
Db      1633 rThrLeuProProThrThrThrProSerProProThrThr-----ThrThrTh 1650
Qy      1238 TCCGTAACGAGCTGCATGCCAAAACCTTCAATTGTGACGCGCAACGATCTTCATCGGC 1297
Db      1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
Qy      1298 TCATTCACTCGACCGCGTTCGCGCGCTCAATACGGAATGGGCGTGTATCGAA 1357
Db      1669 eThrThrThrThrProProProThrThrThrProSerSerProIleThrThrThrPr 1689
Qy      1358 AGCCCAAAATCGCAGACAGATGGAGCGCCCTCGCGCATACACACCCGGAATACGCC 1417
Db      1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
Qy      1418 TACCGGTACCTCGACAAACACAAACCGCTGCAATGGCAGATCCCGCCACCCGAAAA 1477
Db      1708 oile-----ThrThrThrThrThrProSerSerThrThrThrThrProSerProPr 1725
Qy      1478 ACCTACCGCAAGCAACCGGACCAACTTTGGAAACGCATCGCGCAAAATCCTATCC 1537
Db      1725 oThrThrMetThrThrThrProSerProThrThrThrProSerProProThrThrThrMetTh 1745
Qy      1538 CTGCTGCCCA 1547
|||
```

```
Db      1745 rThrLeuPro 1748
RESULT 12
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833.263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1068

Alignment Scores:
Pred. No.: 6,05e-14 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: Gaps: 22

US-10-665-990A-13 (1-1561) x US-09-833-263-1068 (1-5179)
Qy      119 ACGGAAAGCCGTCAATTTCAAT-----ACTTCCAAACCTGCTCTCTCGACAAAC 166
Db      1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
Qy      167 ATCTGCAAAATCCGGCACACCCCTCATAAACACGGGCTATCCGACATCTACTGCTGAC 226
Db      1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
Qy      227 GACCCACAGAACCCCTTGCGCGCGCGCCCTTATCGAATCTGCGAACACAGCCCTC 286
Db      1287 rThrProThr-----SerSerTh 1293
Qy      287 GATTTCGAATACTACTATTGGCGCAACGACATTTCCGCGAGCGTCTGTTCAACCTCA-- 344
Db      1293 rValLeuSerThrThr-----ProlLysLeuCysCysLeuTrpSerAs 1307
Qy      344 ----- 344
Db      1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
Qy      345 -----TGTACCTTGGCGAGAACCGCGCGCTGCGGTACGCC----- 380
Db      1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
Qy      380 ----- 380
Db      1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
Qy      380 ----- 380
Db      1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
Qy      381 -TGCTGTGG-----ACGACAAACAAC 400
Db      1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrThrProSe 1407
Qy      401 ACGCGCGGGTTGGACGATCTCTGCTGCGCTCGACAGCCATCCCAATATCGAAGTGGCG 460
Db      1407 rProProProThrThrThrThrLeuProProProThrThrThrPro----- 1422
|||
```



```
Qy 345 -----TGTACTTGGCGGCAAGCGCGCGTGGCGTACGCC----- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
Qy 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
Qy 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrIleArgValAs 1387
Qy 381 -TGCTGTTGG-----ACGACAACAAC 400
Db 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
Qy 401 ACGCGCGGTGGAGTCTCCTGCTCGCCCTCGACAGCCATCCCATATCGAAGTGGCG 460
Db 1407 rProProProThrThrThrLeuProProThrThrPro----- 1422
Qy 461 CTGTTCAACCCCTTCGTCTACGCAATGGCGCGCACTCGGCTACCTGACCGACTTCCC 520
Db 1423 -----SerProProThrThrThrThrThrProProProThrThrThrProSerProPr 1441
Qy 521 GCGCTCAACCGCGCATGCACAACAATCCTTTACCGCGCAACCGCGCCACCACATACTC 580
Db 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
Qy 581 GCGCGAGCAATATCGGCGGAGCAATACTTCAAGTCGGTGAGGACACCGTTTTCGCGAC 640
Db 1457 -----ProIleSerThrThrThrProProProThrThrThrProSerProProTh 1474
Qy 641 CTGACATCTCGCCACCGGCGAGCTGCTCGCGCAAGTATCGCAGCACTTCGACC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
Qy 696 -----GCTACTGGCAAGCATTCGCGCCACACGCCAGCGCATCATCCGCGCGGC 748
Db 1494 oProProThrThrThrProSerProProMetThrThrThrProIleThrProProIleSerTh 1514
Qy 749 AAGATCGGCAAGGTCTTCAAGCACTCGGATACAAACAGCAAGAAACATCAAGACGCGCTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrThr 1529
Qy 809 CTGCGCTACCGGAAACGTCGAAC-----AGTCGCGCCCTCTACCAAAAAATACACAG 862
Db 1529 rThrThrThrProProProThrThrThrProSerProProThrThr----- 1544
Qy 863 GGAGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGA 922
Db 1545 -----ThrProIleThrProPr 1550
Qy 923 CTCGACCGGACCGCGGAAACCGCGGATTCGCGGAGGTGCAAGACGGCTCAACACAG 982
Db 1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
Qy 983 CCGGAAAAAGCGTCTATCTGGTTTCACCTATTTCGTCCTACAAATCGGCGACAGAC 1042
Db 1563 rProProProThrThrThrThrThrProProProThrThrThrProSerProProThrTh 1583
Qy 1043 GCACTGGCAAAACTGTGTGACGAGCGCATAGACGTTTACCGTCTCTGACCA-----AC 1093
Db 1583 x-----ThrThrProSerProProThrThrThrTh 1593
Qy 1094 TCGCTACAGGGACCGAGGTTCGCGCGTCCATTCGCGTACGTCAAAAT----- 1142
Db 1593 rThrThrProProProThrThrThrProProThrThrThrThrThrThrProPr 1613
Qy 1143 -----ACGGAACCGGTCTCAAGCGCGCATCAAACTC 1177
Db 1613 oProThrThrThrProSerProProThrThrThrThrProIleThrProProThrSerThrTh 1633
```

```
Qy 1178 TACGAGCTGCAACCAACCATGCGTCCCGCCACAAAAGAGGCGCTGACCGGACG 1237
Db 1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
Qy 1238 TCCGTAAACAGCTGATGCAAAACCTTTCATTTGGACGCAAAACGATCTTCATCGGC 1297
Db 1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
Qy 1298 TCATTCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAATGGGGTCTGTCATCGAA 1357
Db 1669 eThrThrThrThrThrThrProProProThr----- 1689
Qy 1358 AGCCCCAAAATCGCAGAACAGATGAGCGACCGACCTCGCCGATACACACCGCAATACGCC 1417
Db 1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerPr 1708
Qy 1418 TACCGCTTACCTCGACAAACACCGCTGCAATGGACGATCCCGCACCGCAAAA 1477
Db 1708 oIle-----ThrThrThrThrProSerSerThrThrThrProSerProPr 1725
Qy 1478 ACCTACCGAAGCAACCGAAGCAAACTTTGGAAACGATCGCGGCAAAATCCTATCC 1537
Db 1725 oThrThrMetThrThrProSerProThrThrThrProSerProProThrThrThrMetTh 1745
Qy 1538 CTGCTGCCCA 1547
Db 1745 rThrLeuPro 1748

RESULT 15
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 1367
; TYPE: PR1
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

Alignment Scores:
Pred. No.: 3,62e-13 Length: 1367
Score: 323.50 Matches: 147
Percent Similarity: 40.7% Conservative: 88
Best Local Similarity: 25.4% Mismatches: 234
Query Match: 11.3% Indels: 109
DB: 3 Gaps: 18

US-10-665-990A-13 (1-1561) x US-09-801-368-108 (1-1367)
Qy 80 TGTTCATGTTCTTCATGTTGCCCACTG-----GAAGAAGCGACG 121
```


; ORGANISM: Homo sapiens

GENERAL INFORMATION:

APPLICANT: Lorens, James B.
 APPLICANT: Atchison, Robert E.
 APPLICANT: Friera, Anabella
 APPLICANT: Holland, Sacha
 APPLICANT: Rigel Pharmaceuticals, Inc.
 TITLE OF INVENTION: Modulators of Angiogenesis and Tumorigenesis
 FILE REFERENCE: 021044-005820US
 CURRENT APPLICATION NUMBER: US/10/696,909A
 CURRENT FILING DATE: 2003-10-29
 PRIOR APPLICATION NUMBER: US 60/512,251
 PRIOR FILING DATE: 2003-10-17
 PRIOR APPLICATION NUMBER: US 60/421,989
 PRIOR FILING DATE: 2002-10-29
 NUMBER OF SEQ ID NOS: 72
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 44
 LENGTH: 2752
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: splicing coactivator subunit SRm300; RNA binding
 OTHER INFORMATION: protein; AT-rich element binding factor
 US-10-696-909A-44

Alignment Scores:

Pred. No.: 8,91e-12 Length: 2752
 Score: 303.50 Matches: 156
 Percent Similarity: 37.6% Conservative: 78
 Best Local Similarity: 25.1% Mismatches: 233
 Query Match: 10.6% Indels: 156
 DB: 5 Gaps: 29

US-10-665-990A-13 (1-1561) x US-10-696-909A-44 (1-2752)

QY 43 AACACGAGCCTCATTTCCCTTTATGCGCTCTCTCTGTTTCATGTTCTTCATGGTGGCC 102
 Db : : : : :
 Db 358 SerThrGlyProGluProProAlaProThrProLeuLeu----- 370
 QY 103 CCCACTGGAAGAACGGAGCGGTCATTTCAATCTTCCAACTGTCCTCTCGGA 162
 Db : : : : :
 Db 371 ---AlaGluArgHisGlyGlySerPro-----GlnProLeuAlaThr 387
 QY 163 CAACATCTCTGCAATCCG-----GCACACCCCTCATAAACAGGGCTATCCGACAT 213
 Db : : : : :
 Db 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProLys 407
 QY 214 CTACTGCTCGACGACCC-----CCACGAAGCCCTTGCGGC 249
 Db : : : : :
 Db 408 SerProGluLysLeuProGlnSerSerSerSerGluSerSerProSerPro---Gln 426
 QY 250 CCGCGCGCCCTTATCGAATCTGCGAACACAGCTCGATTGCAATCTACATTTGGCG 309
 Db : : : : :
 Db 427 ProThrLysValSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla 446
 QY 310 CAACGACATTTCCGCGAGCTGCTTTCAACT----- 342
 Db : : : : :
 Db 447 ProGlySerHisArgGluLeuSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
 QY 343 -----CATGTACTTGGCGCAGAACCGCGGCTGCGGTACG 378
 Db : : : : :
 Db 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
 QY 379 CCTGCTGTGGACGACACACACGCGCGGTGGAGCATCTCCT----- 423
 Db : : : : :
 Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
 QY 424 GCTCGCCCTCGACACCATCCCAATATCGAAGTGGCGCTGTCAACCCCTTCGTCCTACG 483
 Db : : : : :
 Db 506 SerArgSerArgSerProGlnThrArgArgSerArgSerAlaGln----- 520
 QY 484 CAAATGGCGGCATCGGCTACCTGACCGCATTCCTCCCGCGCTCAACCGCGCATGCAAA 543
 Db : : : : :

Db 521 ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro----- 537
 QY 544 CAAATCCTTTTACCCCGGACACACCGCGCCACCATCTCGCGGAGCAGCAATATCGGCGACA 603
 Db : : : : :
 Db 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
 QY 604 ATACTTCAAAGTCGGTGAGGACACCGTTTTCGCGACCTGGACATCTCTCGCCACCGCAG 663
 Db : : : : :
 Db 554 SerArgSerAlaArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
 QY 664 CGTCGTGCGGCAAGTATCGCACGACTTCGACCGCTACTTGGCAAGCCATTCGCCCAACA 723
 Db : : : : :
 Db 574 ArgSerArgThrProAlaArgArgGlyArgSerArgThrProAlaArgArg 593
 QY 724 CGCCACGGCATCATCCGACGCGCAACAT-----CGCAAGG 762
 Db : : : : :
 Db 594 SerArgSerArgThrProThrArgArgSerArgSerArgThrProAlaArgGly 613
 QY 763 TCTTCAAGCACTCGGATA---CAACGACGAAACATCCAGACACGCGCTCTCGCTACCG 819
 Db : : : : :
 Db 614 ArgSerArgSerArgThrProAlaArgArgSerArgThrArgSerPro---ValArg 632
 QY 820 CGAAACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGACGGGCGCATCGATGCA 879
 Db : : : : :
 Db 633 ArgArgSerArgSerProAlaArgSerGlyArgSerArgSerArgThrPro 652
 QY 880 GAGCGTCCAAACCGCCTGATCAGCAGACACCCCTGCAAA-----AGCACTCGACCGCA 933
 Db : : : : :
 Db 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
 QY 934 CCGCGCAAAACCGCGGATTCGCGGAGCTGCAAGA----- 969
 Db : : : : :
 Db 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
 QY 970 CGCGCTCAACAGCCGCAAAAGCGTCTATCTGGTTTCACTTTCCTTTCCTTACAAA 1029
 Db : : : : :
 Db 693 ArgSerArgThrProArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg 712
 QY 1030 ATCCGGCACAGA-----CGCACTGGCAAAACTGGTCAGACGCGCATAGCGTTAC 1080
 Db : : : : :
 Db 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerGluArgLys--- 731
 QY 1081 CGTCCTCACCAACTCGCTACAGCGACCGAGCTGTCGCGC----- 1119
 Db : : : : :
 Db 732 -----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet 749
 QY 1120 -----CGTTCATTCGGCTACGTCAAAATACCGAAACCGCTGCT 1158
 Db : : : : :
 Db 750 LysLysSerArgIleSerSerArgSerArgSerLeuSerSerProArgSerLysAla 769
 QY 1159 CAAAGCCGGCATCAAACTCTACGA-----GCTGCAACCCCAACCATCGCTCCCGCCAC 1212
 Db : : : : :
 Db 770 LysSerArgLeuSerLeuArgSerLeuSerGlySerProLysProLysGln 788
 QY 1213 AAAAGACAAAAGCGCTGACCGCGCTCGT-----AACCAAGCTGC----- 1253
 Db : : : : :
 Db 789 LysSerGlnThrProProArgArgSerArgSerGlySerGlnProLysAlaLysSer 808
 QY 1254 -----ATGCCAAACCTTCATTGTGGACGCGCA 1280
 Db : : : : :
 Db 809 ArgThrProProArgArgSerArgSerSerSerProProLysGlnLysSerLys 828
 QY 1281 AACGCATCTTCATCGGCTCATCAACCTCGACCCCGCTTCG----- 1322
 Db : : : : :
 Db 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
 QY 1323 -----CACGCTCAATACCGAAATGGCGCTCGTCATCGAAAGCCCAAAATTCGAG 1373
 Db : : : : :
 Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluGlnSerValThrProGln 868
 QY 1374 AACAGATGGAGCGACCCCTCGCGATACACACCGGAATACGCTACCGCGTTACCTCG 1433
 Db : : : : :
 Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887

Db	789	LysSerGlnThrProProArgSerArgSerGlySerSerGlnThrLysAlaLysSer	808
Qy	1248	-----GCCTGCATGCCAAACCTTTCATTTGGACGCCA	1380
Db	809	ArgThrProProArgSerArgSerSerProProLysGlnLysSerLys	828
Qy	1381	AACGCATCTTCATCGGCTCATTTCAACCTCGACCCCGTTCCG	1322
Db	829	ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr	848
Qy	1323	-----CACGGCTCAATACCCGAAATGGCGTCGTCATCGAAGCCCAAAATCGCAG	1373
Db	849	ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluGlnSerValThrProGln	868
Qy	1374	AACAGATGGAGCGCACCTCGCCGATACACACCCGAATACGCCTACCGCGTTTACCCCTCG	1433
Db	869	ArgArg--SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer	887
Qy	1434	ACAACACACACCGCTCGCATGGC-----AGATCCCGCCCA	1469
Db	888	ArgHis-----SerCysSerGlySerSerProProArgVallysSerSerThrProPro	905
Qy	1470	CCCGAAAAACCTACCCCGA--ACGAACCCGAAGCCA-----	1502
Db	906	ArgGlnSerProSerArgSerSerSerProGlnProLysVallysAlaIleSerPro	925
Qy	1503	-----AACTTTGGAACGCATCGCGCGCAAAATCCTATCCTGCTGCCCATCG	1550
Db	926	ArgGlnArgSerHisSerGlySerSerProSer	939
Qy	1551	AAGGTT 1556	
Db	940	ArgVal 941	
RESULT 20			
US-09-840-746-20			
; Sequence 20, Application US/09840746			
; Publication No. US20030166501A1			
; GENERAL INFORMATION:			
; APPLICANT: Chen, Hwei-Mei			
; APPLICANT: Honchell, Cynthia D.			
; APPLICANT: Tang, Y. Tom			
; TITLE OF INVENTION: Mucin-Related Tumor Marker			
; FILE REFERENCE: PC-0039 US			
; CURRENT APPLICATION NUMBER: US/09/840,746			
; CURRENT FILING DATE: 2001-04-23			
; NUMBER OF SEQ ID NOS: 20			
; SOFTWARE: PERL Program			
; SEQ ID NO 20			
; LENGTH: 528			
; TYPE: PRT			
; ORGANISM: Sus scrofa			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Genbank ID No. US20030166501A1 g915208			
US-09-840-746-20			
Alignment Scores:			
Pred. No.: 1,17e-11 Length: 528			
Score: 300.00 Matches: 137			
Percent Similarity: 41.3% Conservative: 92			
Best Local Similarity: 24.7% Mismatches: 179			
Query Match: 10.5% Indels: 146			
DB: 3 Gaps: 24			
US-10-665-990A-13 (1-1561) x US-09-840-746-20 (1-528)			
Qy	132	ATTTCATACTTCAAACCTGTCTCTCC-----TCGACACATCCTCGAAA	176
Db	2	IleSerValGlnProSerSerSerSerProThrThrThrSerValGln	21
Qy	177	TCGGGCACA-----CCCTCATACACAGGGCTATCCGACATCT---	215

```
QY 1146 -----GAAACCGTGCTCAAGCCGGCATCA 1172
Db 348 ProProfileSerSerThrIleSerValGlnProSerSerSerSerProThrThr 367
QY 1173 AACTCTAGAGGTGCAACCAACCATCGCGTCCGCCACAAAGACAAAGCCGTGACCG 1232
Db 368 SerThrThrSerValGlnProSerSerSerGlySerAlaProThrThrSerAlaThrSer 387
QY 1233 GCAGCTCCGTAAACCACTGTCATGCCAAACCTTCATTGTGACGGCAAAACGCATCTTCA 1292
Db 388 ValGlnPro-----SerSerSerSerSerValProThrThrSerAlaThrSer 403
QY 1293 -----TGGGCTCATTAACCTGACCCCGTTCGCGACGGCTCAATACCGAAATGGGG 1346
Db 404 ValArgSerSerSerSerThrProIleProThrThrThrSerValGlnProSer 423
QY 1347 TCGTCATCGAAGCCCAAAATCG-----CAGAACAGATGGAGCGCACCC 1391
Db 424 SerSerSerSerValProThrThrSerAlaThrSerValGlnThrSerSerSerSer 443
QY 1392 TCGCGGATACCA-----CACCGGAATACGCTACCGGTTTACCTCGACA 1436
Db 444 ThrProIleProSerThrThrSerValGlnProSerSerSerSerSerAlaProThrThr 463
QY 1437 AACACAAACCGCTGC-----AATGGCAGATCCCGCCACCGAAGAAACCTTACCGAAG 1490
Db 464 SerAlaThrSerValGlnProSerSerSerSerProIleSerSerThr---Ile 482
QY 1491 AACCCGAAGCAAACTTTGGAACGATCGCGCAAAATCC 1532
Db 483 SerValGlnProSerSerSerSerSerProThrThrSer 496

RESULT 21
US-10-437-963-125004
; Sequence 125004, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125004
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(497)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27690C.1.pap
US-10-437-963-125004

Alignment Scores:
Pred. No.: 1.57e-11 Length: 497
Score: 298.00 Matches: 145
Percent Similarity: 37.7% Conservative: 56
Best Local Similarity: 27.2% Mismatches: 170
Query Match: 10.4% Indels: 162
DB: 4 Gaps: 29

US-10-665-990A-13 (1-1561) x US-10-437-963-125004 (1-497)
```

```
QY 267 AATCTGCCGAACACAGCCTCGATTTC-----AATACTACA 302
Db 1 AsnLeuHisAsnAsn---AsnLeuCysGlySerValProValArgGluGlyAsnIysThr 19
QY 303 TTGGCGCAACGACATTTCCGCGAGCG-----TGCTGT-----TCAACC 341
Db 20 GlyProSerSerSerGluValAlaGlySerSerCysCysAspGluAspGluAlaAspLeu 39
QY 342 TCATGTACTTTCGCGCAGAACGCGGTACGCTGCTGCTGTGTGGACGCAACAACA 401
Db 40 SerCysLeuMetProThrAlaAlaArgAlaSerAlaAlaAala---AlaValThrAla 58
QY 402 CGCGGGGTGGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAG----- 455
Db 59 ArgSerSerSerSerSerAlaGlyProAlaThrSerAlaThrSerSerGlyArg 78
QY 456 -----TGCGCTGTTCACCCCTTCGTCCTAC 482
Db 79 SerArgGlyAlaProProAlaSerTrpArgProAlaProCysSerThrProProSer*** 98
QY 483 GCAATGGCGGCATCTCGGTACTCTGACCGACTTCCCGCGCTCAACCGCGCATGCACA 542
Db 99 SerAlaAlaProTrpArgAla-----ProThrThrProThrSerAlaArgSerThr 116
QY 543 ACA-----AATCTTTACCGCCGACACCGCGCCA 572
Db 117 SerArgArgProAlaProSerArgProArgGlySerProProProProThrSerArg 136
QY 573 CCATACTCGCGGACGCA-----ATATCGCGGACG 602
Db 137 ProThrSerSerSerThrArgSerSerArgAlaSerProSerProProSerAlaSer 156
QY 603 AATACTCAAAGTCGTGAGGACACCGTTTCGCGCAGCTGGACATCTCGCCACCGCA 662
Db 157 AlaProSerProAlaThrGlyThrProProSerProProThrThrSerSerProThr 176
QY 663 GCG-----TCGTGCGCGAAGTATCGCAGCT 689
Db 177 AlaProGlyProProProArgAlaThrArgSerCysSerSerSerArgArgThrThr 196
QY 690 TCGACCGCTACTGGGCAAGCCATTCGCGCCACACGCGCAGCATCTCCCGCAGCGCA 749
Db 197 AlaThrPro-----ProProSerThrProAlaArgSerAlaAlaAlaSer 211
QY 750 ACATCGGCAAGGTCTCAAGCAGCTCGGATACAACAGCAACATCCAGACAGCGCTCC 809
Db 212 ArgArgProProHisAlaSerSer***ProSerThrThrSerAlaProSerThrProSer 231
QY 810 TGC-----GCTACCGCGAAACCGTCGACAGTCGCGCCCTCTACCAAAAAA 854
Db 232 CysArgProArgAlaGlyAlaSerProProAlaGlyAlaProAlaProGlyGlu 251
QY 855 TACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCGACACCCCTG 914
Db 252 CysArgArg-----ProArgProAla---SerSerGlnProCys 263
QY 915 CAAAAG---GACTCGACCGGACCGCCGCAACCGCGATTCCGGGAGGCTGCAAGACG 971
Db 264 ArgArgCysSerSerThrAlaThrSerThr-----GlyCys----- 275
QY 972 CGCTCAACAGCCCGAAAAAGCGTCTATCTGTTTCCACCTTATTTCTGTCCTTACAAAT 1031
Db 276 ---SerSerGlnProProProAlaThr-----SerArgSerGlyGluSerThrArgSer 292
QY 1032 CGGCAACAGACGCACTGGCAAACTGGTGCAGAGCGCATAGACGTTACCG-----TCC 1085
Db 293 ProSerAlaPro-----SerSerSerGlyArgSerLeuProThrThrSer 307
QY 1086 TGACCAACTCGCTACAGGCGA-----CGACGTTGCGCGCTCCATTCGCGTACGTCA 1139
Db 308 SerSerThrArgSerSerArgAlaSerProProProProSerAspSerAlaProSer 327
```



```
QY 1140 AATACCGAAACCGCTGCTCAAAGCGGCATCAAACTCTACGAGCTGCAACCAACCAACCATG 1199
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
328 ProAlaThrGly-----ThrProSerProProThrThrSerSerSerProThrAla 345
QY 1200 CGTCCCGCCA-----CAAAGCAAAAG 1223
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 ProGlyProProProAlaAlaThrArgSerPheSerSerSerArgArgArgAsnPro 365
QY 1224 GCCTGACCGGAGCTCCGTAACAGCCTGCATGCGCAAAACCTTCATTGTCGACGCAAAAC 1283
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
366 AlaGluProAlaAlaProPro-----SerThr 375
QY 1284 GAATCTCATCGCTCATTCACCTCGACCCCGCTTCGACGGCTCAATACCGAAATGG 1343
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
376 ProAlaArgSerAlaAlaAlaArgArgProProProAlaSerSerSerProSerThr 395
QY 1344 CGCTGCTCATCGAAAGCCCCAAATCGCAGAACAGATGAGGGCACCTCG---CCGATA 1400
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 Ala-----SerAlaProSerThrProSer 403
QY 1401 CCACACCCGAATACGCTACCGCTTACCTCGACAAACACACCGCTGCAATGGC--- 1457
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
404 ArgArgProGlyArgAlaGlyAlaSerProSerSerProThrAlaAlaSerArgGlyIle 423
QY 1458 -----ACGATCCCGCCA-----CCGAAACCACTACC----- 1484
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
424 ThrSerSerThrCysProProAlaIleThrSerArgProThrSerSerProAlaCysHis 443
QY 1485 -----CGAACGAACCCGAACCGAACCTTTGGA 1511
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
444 ArgCysSerSerThrAlaThrSerThrGlyCysSerSerGlnProProProAlaThrGly 463
QY 1512 AACGCGATCGCGCAAAATCCTATCCCTGCTGCCATCG 1550
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
464 ArgSerGlyGlySerCysArgSerProSerAlaProSer 476
```

RESULT 22

```
US-10-084-846A-5
; Sequence 5, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREFFZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 5
; LENGTH: 19723
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-5
```

Alignment Scores:

Pred. No.:	3.11e-11	Length:	19723
Score:	297.50	Matches:	158
Percent Similarity:	37.0%	Conservative:	63
Best Local Similarity:	26.5%	Mismatches:	212
Query Match:	10.4%	Indels:	165
DB:	4	Gaps:	29

US-10-665-990A-13 (1-1561) x US-10-084-846A-5 (1-19723)

```
QY 100 GCCCCCACTGGAAGAACGGACGGAAAGCCGTCATTTCAATACTTCCAAACCTGTCTCTCT 159
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18236 AlaGlyThrGlyArgSer-----ProAla 18243
QY 160 GGACAAACATCTGCAAAATCCGGCACACCCCTCATAACACACGGCTATCCGACATCTACT 219
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18244 CysSerArgProArgGlnProThrCysProSer---CysArgGlyValArgProAlaPro 18262
QY 220 GCTCGA-----CGACCCCGACGAAGCCCTTGCCGC-----CCGCGC 255
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18263 GlyArgValArgAlaAlaGlyAlaAlaSerProValArgSerThrArgTrpSerProArg 18282
QY 256 CGCCCTTATCAAACTCGCG---AACACACCGCTCG-----ATTGCAATACTACATT 305
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18283 ArgArgGlyProLeuCysProGlyAsnSerAlaSerProProCysGlyArg----- 18300
QY 306 GCGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGT-----ACCTTGCCGCAG 359
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18301 -----ProAlaSerAlaCysProProSerAlaTrpSerThrGlySerSer 18315
QY 360 AACCGCGCTGCGGTACGCTGCTGTGGACGACAAACACACCGCGGTGGACGATC 419
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18316 SerAlaAlaAlaSerTrpSerCysTrpProAsnSerProHisAlaProIleArgThr 18335
QY 420 TCCTGC-----TCGCCCTCGACACGCCATCCCAATA-----TCGAAG 455
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18336 ArgCysArgProValGlnAsnProSerThrGlyAlaArgThrAspLysSerGlnSerArg 18355
QY 456 TGCCTCTGTTCAACCCCTTCGCTACGCAATGGCGCCTCGGCTACCTGACCGGACT 515
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18356 CysAlaThrPheValProThrThrGlnThrAsnSer-----ProThr 18369
QY 516 TCCCCCGCTCAACCGCGCATGCACAAACAAATCCTTTACCG---CCGACAAACCGCGCA 572
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18370 SerAlaThrSerSerThrHisAlaThrAlaAspProSerProLeuProAlaThrPro 18389
QY 573 CCATACCTCGCGGACGCAATATCGCGACGCAATACT-----TCAAGTCG 617
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18390 GlyHisSerThrGlyAlaSerProThrGlyThrGlyProThrThrTyrSerSerAla 18409
QY 618 GTGAGGACACCGTTTTCGCGACCTGCACATCTCTCGCACCGCA----- 662
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18410 AlaSerTrpArgProThrProSerSerThrAlaSerArgProAlaMetValSerPhePhe 18429
QY 663 -----GCGTCGTCGCGCAAGTATCGCACAGCTTCGACCGCTACTGGGCAA 707
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18430 AlaSerAlaMetThrAlaThrSerSerAlaSerArgSerThrThrAlaGluArgAlaTyr 18449
QY 708 GCCATTCCGCGCCCAACGCCA---CGCGCATCATCCGACGGGCAACATCGGCAAGGTC 764
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18450 ProGluProProThrThrProThrLysAlaAlaAlaValCysSerTrpSerAlaArg--- 18468
QY 765 TTCAAGCACTCGGATACACACGCAACATCCAGACACGCGCTCTCGCTACCGCGAAA 824
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18469 -----TrpProThrSerGlyAlaTrpAlaSerGluArg 18479
QY 825 CGGTGCAACAGTCGCCCTCTACCAAAAAATACAGA-----CGG 863
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18480 ProAlaLysSerSerGlyProSerSerProCysArgAlaLeuValGlnPheLeuTrpArg 18499
QY 864 GACGCATCGACTGCGACAGCGT-----CCAAACCCGCC 896
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18500 ProGlnSerPheGlyArg-ThrGlyGluAlaGlnValAspMetProLeuProThrProPr 18519
QY 897 TGATCAGCGACACCCCTGCAAAAGGACTCGA----- 927
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18519 o-----ArgArgProHisLysArgThrThrAspGlyGlyProLeuArgTrpGlnCysVa 18537
QY 927 ----- 927
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18537 lIleProGlyGluAspTrpProProSerAlaArgProProArgProSerProArgAlase 18557
```

```
QY 928 -----CCGCGACCGCGCAAAACCCGCGATTGCCGGGAGGCTGCAAGACG 971
Db 18557 rSerProLeuSerProProArgProProArgAlaValSerSerThrArgAlaArgG1 18577
QY 972 C-----GCTCAAAACAGCCGAAAAAGCGTCTATCTGTTTCCACCCTATTTCG 1019
Db 18577 nThrProTrpProThrAlaThrSerSerArgArgThrThrProValArgValProAr 18597
QY 1020 TCCCTACAAAATCCGGCAGACGACCTGGCAAAACTGGTCAGCAGCGCATAGACGTTA 1079
Db 18597 gProAspAlaProSerSerArgSerThrAlaProThrSerSerAlaAaGThrArgArg-- 18616
QY 1080 CCGTCCTGACCAACTCGCTACAGGCGACGCGAGTTCGCCCGCTCATCTCCGGCTACGTCA 1139
Db 18617 -----ProThrAlaThrArgSerArgProProArgArg-----ArgProArgAr 18631
QY 1140 AATACCGAAAAACCGTGCTCAAAAGCGGCATCAAACTCTACGAGCTGCAAC----- 1190
Db 18631 gSerProProThrArg--ArgSerProProTrpSerArgThrAlaCysSerAlaSerThr 18650
QY 1191 -----CCAACCATGCCG-----TCCCGG 1208
Db 18651 ArgProArgProThrArgProProGlyAlaTTPThrAlaTTPThrArgArgThrSerPro 18670
QY 1209 CCACAAAAGACAAAGCCCTGACCGCAGCTCGTAACGAGCTGATGCCAAAACCTTCA 1268
Db 18671 ThrAlaArgThr-----ProThrArgThrArgProGlyArgGluArgArgThrSer 18687
QY 1269 TTGTGGCGGCAACGATCTTCATCGGCTCATCAACCTCGACCCCGTTCGCGCAGCGC 1328
Db 18688 SerThrProAlaSerAlaSerProThrAlaThrSerAlaAlaGlyPro---ProThrAla 18706
QY 1329 TCAATACCGAAATGGCGCTGTCATCGAAGGCCCAAAATCCGACAGACAGATGGAGCGCA 1388
Db 18707 ThrThrPro-----SerThrThrThrProProArgThrAlaThrAlaThrAlaArg 18724
QY 1389 CCCTCGCGATACACACCGCAATACGCTACCGGTTACCTCGACAAACACACACCGCC 1448
Db 18725 ThrTrpProAlaArgSerProAlaThrProThrAlaSerProArgProArgSerAla 18744
QY 1449 TGCAATGGCAGCATCCCGCCACCCGAAACAACTTACCCGAAACCCGAAAGCCAAACTTT 1508
Db 18745 SerAlaCysThrThrProAlaArgAlaProProProArgSerSerProAlaSerThrGly 18764
QY 1509 GGAACGCGATCCGCG---CAAAAAATCTATCTCTCTGCTGCCCATCG 1550
Db 18765 SerProGlyThrProSerSerArgProSerProThrCysProSer 18779
RESULT 23
US-10-418-861B-59
; Sequence 59, Application US/1041861B
; Publication No. US20040010131A1
; GENERAL INFORMATION:
; APPLICANT: da Silva, Ana Claudia Rasera
; APPLICANT: Farah, Shaker Chuck
; APPLICANT: Quaggio, Ronaldo Bento
; APPLICANT: Reinach, Fernando de Castro
; APPLICANT: Ferro, Jesus Aparecido
; APPLICANT: De Oliveira, Julio Cezar Franco
; APPLICANT: De Laia, Marcelo Luiz
; APPLICANT: Setubal Joao C.
; APPLICANT: Furlan, Luiz Roberto
; TITLE OF INVENTION: Isolated Xanthomonas nucleic acid molecules, proteins encoded the
; FILE OF INVENTION: uses thereof
; FILE REFERENCE: FAPESP 205.1 US
; CURRENT APPLICATION NUMBER: US/10/418,861B
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/374,620
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 59
; LENGTH: 486
; TYPE: PRT
```

```
; ORGANISM: Xanthomonas
; FEATURE:
US-10-418-861B-59
Alignment Scores: 2,13e-11 Length: 486
Score: 296.00 Matches: 123
Percent Similarity: 38.0% Conservatives: 63
Best Local Similarity: 25.2% Mismatches: 209
Query Match: 10.4% Indels: 94
DB: 4 Gaps: 14
US-10-665-990A-13 (1-1561) x US-10-418-861B-59 (1-486)
QY 83 TCATGTTCTTCATGGTTGCCCCCACTGGAAGAACGGACGGAAGCCGTCATTTCATTA 142
Db 89 SerTyrSerValCysProProAspAlaAspCysThrGlu----- 102
QY 143 TCCAAACCTGTCTCCTCGGACAAACATCTGCAAAATCCGGCACACCCCTCATACAACACGG 202
Db 103 -----LeuAlaLysIleAlaGlnSerThrThrGlyLeuAlaProSerSer 117
QY 203 CTATCCGACATCTACCTGCTCGACGACCCCCACGAAGCCCTTGGCGCGCGCCGCTT 262
Db 118 AlaThrGluValHisTrpLeuValAspGlyAlaAlaThrTyrAlaAlaIleIleGluAla 137
QY 263 ATCGAATCTGCGCAACACAGCCTCGATTGCAATCTACTACTATTGGCGCAACGACATTTCC 322
Db 138 IleArgGlyAlaAaGAspHisIleHisLeuGluTyrTyrIlePheGlnProAspHisSer 157
QY 323 GGCAGGCTGCTGTCAACCTCATGTACCTTCCCGCAGAACCGGGCGTGGCTGACGCTG 382
Db 158 GlyThrAlaIleCysAlaAlaLeuMetGluArgAlaGlyValLysValArgLeu 177
QY 383 CTGTGGACGACAAACACACGCGGGTTGACGACATCTCTGCTCGCCCTCGACACCAT 442
Db 178 LeuMetAspAlaIleGlySerSerAlaMetThrArgArgAlaLeuArgThrLeuArgGlu 197
QY 443 CCCAATATCGAAGTGGCGCTTTCACCCCTTGTCTTACGCAAAATGGCGGCGCATCGGC 502
Db 198 AlaGlyValGluThrAlaTrpPheHisProSerGlnLeu----- 210
QY 503 TACTGTACCGACTTC-----CCCGCCTCAACCGCGCATGTCACAAACATCTTTACC 556
Db 211 ---LeuLysProPheLysArgProTrpLeuAsnLeuArgThrHisArgLysValIleVal 229
QY 557 GCCGACAAACCGCGCCACCATATCTCGCGGACGCAATATCGGCGAGCAATACTTCAAAGTC 616
Db 230 IleAspGlyArgIleGlyPheThrGlyGlyIleAsnValThrAspAspGluAsnGluGln 249
QY 617 GGTGAGGACACCGTTTCCCGACCTGGACATCTCTCGCCACCGGCGAGCGTGTGCGCGAA 676
Db 250 ValArgLysGlnAlaTyrArgAspLeuHisValArgLeuGlnGlyHisValValArgSer 269
QY 677 GTATCGCAGCACTTCGACCGCTACTTGGGCAAGCCATCTCCGCCCAACCGCACCGCGCATC 736
Db 270 LeuGlnLeuValPheLeuGluAspTrpLeu-----TyrAlaThrSerGln 284
QY 737 ATCCGACGCGGCAACATCGCAAGGGTCTTCAAGCACTCGGATACAAACGACGAAATATCC 796
Db 285 GlyArgAlaAlaPheHisGlyGlnLeu-----TrpProAspAspMetPro 300
QY 797 AGACACGCGCTCTGCGGTACCGGAAACCGTCGNAACAGTCCGCTCTACCAAAAAATA 856
Db 301 ThrArgAla-----GlnGlyThrValAspAlaGlnValLeu-----Val 313
QY 857 CAGACGGACGATCGACTGCGCAGAGCGTCAAAACCCGCTGATCAGGCGACACCCCTGCA 916
Db 314 SerGlyProAspSerSerTrpGluAlaIleHisArgLeuMetValAla----- 329
QY 917 AAAGGACTCGACCGCGACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGACGCGTTC 976
Db 330 -----AlaIle 331
```

Qy	977	AAACAGCCCGAAAAAAGCGTCTATCTGGTTTCACCCATTATTGCTCCCTACAAAATCCGGC	1036
Db	332	HisGluAlaYHisArgValTrpLeuValTrpProTyrPheValProGlyGluAlaAa	351
Qy	1037	ACAGACGACATGCGCAAACTGGTCAGGACGGCATAGACGTTACCGTCTCAGCAACTCG	1096
Db	352	ArgMetAlaLeuThrSerAlaAlaLeuGlyGlyLeuAspValArgLeuLeuValPro	370
Qy	1097	CTACAGCGCAGCGACGTTGGCCGCGCTCCATTCCGGCTACGTCAAATACCGAAAACCGCTG	1156
Db	371	--ArgValSerAspSerArgLeuValThrTyrAlaAlaArgSerTyrPheAspGluLeu	389
Qy	1157	CTCAAGCGCGCATCAAACTCTACGAGGTGCACCCAAACCATGCGCTCCCGCCACAAA	1216
Db	390	LeuGluAlaGlyValArgGlyLeuTyrGlyArg	402
Qy	1217	GACAAAGCGCTGACCGGAGCTCGCTAACAGCGCTGCATGCCAAAACCTTCATTGTGGAC	1276
Db	403	-----MetLeuHisThrLysAlaLeuLeuAlaAsp	412
Qy	1277	GGCAAAAGCGCATCTTCATCGGCTCATTCAAACCTCCAGCCCGCTCCGCAGCGCTCAATACC	1336
Db	413	AspAspValCysIleValGlySerAlaAsnPheAspSerArgSerPheArgLeuAsnPhe	432
Qy	1337	GAATGGGCTGCTCATCGAAGCCCAAAATCCAGAACAGATGGAGCGACCTCGCC	1396
Db	433	GluLeuSerMetLeuPheArgAspGlnAlaValAlaAlaGluMet--AlaGlyLeuIle	451
Qy	1397	GATACCACACCCGAATACGCTACCGGTTACCTCCGACAAACACACACCGCTGCATGG	1456
Db	452	GlyThrAspLeuGlnAlaGlnGluValGlnPheAlaArgHisArgProLeu--Trp	470
Qy	1457	CAGATCCCGCCACCCGAAAAAACCTACCCGAAACGACCGAAGCCAAACTTTGGAAACGC	1516
Db	471	ArgSer-----ArgLeuProGluAla-----	477
Qy	1517	ATCGCCGCAAAAATCCTATCCCTGCTG	1543
Db	478	---PheAlaArgLeuLeuSerProLeu	485

RESULT 24

```

US-10-626-832-86
; Sequence 86, Application US/10626832
; Publication No. US20050003342A1
; GENERAL INFORMATION:
; APPLICANT: Davis Poynter, Nick
; APPLICANT: Nugent, Josephine
; APPLICANT: Birch-Machin, Ian
; APPLICANT: Allen, George P
; TITLE OF INVENTION: Viral Marker
; FILE REFERENCE: 620-262
; CURRENT APPLICATION NUMBER: US/10/6
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/398
; PRIOR FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Equine herpesvirus 1
US-10-626-832-86

```

Alignment Scores:		
Pred. No.:	3.45e-11	Length:
Score:	293.50	Matches:
Percent Similarity:	38.8%	Conservative:
Best Local Similarity:	25.1%	Mismatches:
Query Match:	10.3%	Indels:
DB:	5	Gaps:
		DB:

US-10-665-990A-13 (1-1561) x US-10-626-832-86 (1-866)

Qy	126	GCGGTCAATTTCAATACTTCCAAACGCTGCTCCTGGAGCAACAATCTCTGCAAAATCCGGCAC	189
Db	19	AlaIleGlySerThrThrThrGluThrThrThrSerSerSerSerThrThrSerGlySer	38
Qy	186	CCCCTATAACAACGGGCTATCCGACATCTACTGTCGACGACGCCCCACGAAGCCCTTG	245
Db	39	GlyGlnSerThrSerSerGlyThrThrAsnSerSerSerSerProThrThrSerPro--	57
Qy	246	CCGCCCGCGCCCTTATCGAANTCGCGAACACAGCCTCGATTGGCAATACTACATTT	305
Db	58	-----ProThrThrSerSerSerProProThrThrHisThrSerSerProSer	74
Qy	306	GGCGCAACGACATTTCCGGCAGGCTGCTGTTTCAAACCTCATCTCCGCGCAGAACGCG	365
Db	75	SerThrSerThrGlnSerSerSerThrAlaAlaThrSerSerSerAlaProSerThrAla	94
Qy	366	GCGTGGCGGTACGGCTGCTGTTGGACGACAACAACACGCGCGGGTTGGACGATCTCTGC	425
Db	94	-----	94
Qy	426	TGCGCCTCGACAGCATCCCCAATATCGAAGTGCCTGTTCAACCCCTTCTGCTCTACGCA	485
Db	95	SerSerThrThrSerIleProThrSer-----ThrSerThrGluThrThrThrThr	111
Qy	486	AATGGCGCGCACTCGGTACTCGTAGCG-----ACTTCCCGCGCTCAACCGCGCGCATCG	539
Db	112	ThrProThrAlaSerThrThrThrProThrThrThrAlaAlaProThrThrAlaAla	131
Qy	540	ACAACAATCTTTACCGCGACAAACCGCGCCACCATCTCGGCGGAGCAATATCGGCG	599
Db	132	ThrThrThrAlaVal-----ThrThrAlaAlaSerThrSerAlaGluThrThrThrAla	149
Qy	600	ACGAATACTTCAAAGTCGGTGAGACACCCTTTTCGCGGACCTGGACATCTCTGCCACCG	659
Db	150	ThrAlaThrAlaThrSerThrProThrThrThrThrProThrSerThrThrThrThrThr	169
Qy	660	GCAGCGTCGTCGGCGAAGTATCGCACACTTCGACCGCTACTGGGCAAGCCATTCGCGCC	719
Db	170	AlaThrThrThrValPro--ThrThrAlaSerThrThrThrAspThrThrThrAlaAla	188
Qy	720	ACAAGCCACCGCATCATCGCAGCGGCAACATCGGCAAGGTCCTTCAAGCACTCGGAT	779
Db	189	ThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAla	208
Qy	780	ACAAGCAGAAACATCCAGACACGCGCTCTCGCTACCGCGAAACCGTCGAAACAGTCGC	839
Db	209	ThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAla	228
Qy	840	CCCTCTACCAAAAAATACAGACGGGACGCATCGACTGGCAGAGCGTCCAAACCGCGCC	896
Db	229	ThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrSerSerAlaThrThrAlaAlaThr	248
Qy	897	TGATCAGCAGACCCCTGCAAAAAGGATCGACCGCGACCGCGCAAAACCGCGGATTCGCG	956
Db	249	ThrSerSerThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThr	268
Qy	957	GGAGGTGCAAGACGCGCTCAACACAGCCCGGAAAAAAGGCTATCTGGTTTCACCCCTATT	1016
Db	269	ThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThr	288
Qy	1017	TCGTCCCTACAAAATCCGGCACAGACGCACTGGCAAAACTGGTCAGACGCGCATAGACG	1076
Db	289	ThrAlaAlaThrThrThrAlaAlaThr-----ThrThrAlaAlaThr	302
Qy	1077	TTACCGTCTGACCAACTCGCTACAGCGCACCGACGTTGCGCGCTCATTCCTCGCGTACG	1136
Db	303	ThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThr	322
Qy	1137	TCAAAATACCGAAAACCGCTGCTMAAGCCGCGCATCAAACTCTACGAGCTGCAACCCAACC	1196
Db	323	ThrThrAlaAlaThrThrThrGlySerProThrThrSerGlySerThrSerThrThrGlyAla	342

Qy	1077	TTACCGTCCTGACCAATCGCTCAGCGCGACGACGCTGGCGCGTCCATTCGGGGTCAGC	1136
Db	289	ThrAlaAlaThrThrThrAlaAlaThr-----ThrThrAlaAlaThr	303
Db	303	ThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThr	322
Qy	1137	TCAAAATACCGAAACCGCTGCTCAAAACGGCGCATCAAACCTACGAGGTGCAACCCAACC	1196
Db	323	ThrThrAlaAlaThrThrThrGlySerProThrSerGlySerThrSerThrThrGlyAla	342


```

Db      322 rThrIasnSerProArgSerThrSerValGlyArgArgSerProSerProCysProSerLe 342
Qy      1127 TCGCGCTACGTCAAAATACCGAAACCGCTGCTCAAAAGCGGCATCAAACTCTACGAGCTG 1186
Db      342 u-----SerSerProCysSerSerProProSerProSer----- 353
Qy      1187 CAACCCAACCATGCCGTCCCGCCGACAAAGACAAAGAGCGCTGACCGGACGCTCGTAAAC 1246
Db      354 -----ProSerProSerGlnSerThrProSerSerSerAlaProSerPr 369
Qy      1247 AGCCTGCATGCGCAAAACCTTCATTGTGGACGGCAAAACGATCTTCATCGGCTCATTTCAAC 1306
Db      369 o-----**AlaSerAlaSerAlaSerThrLysSerAl 380
Qy      1307 CTCGACCCCGTTCCGCACGGCTCAATACCGAAATGGCGTCGTCATCGAAAGCCCAAA 1366
Db      380 aSerValValIleArgGly-----ValSerSerAlaProSe 393
Qy      1367 ATCGCAGAACATGAGGGGCAACCTCGCGGATACACACCCGAATACGCCT----- 1418
Db      393 r-----SerAlaProSerProProProSerSerSerProProSe 408
Qy      1419 -----ACCGCGTTACCTCGACAAACACACACCGCCTGCAATGGCAC 1459
Db      408 rSerSerAlaArgArgArgSerProSerProThrArgArgTptThrSerCysSerGlyTh 428
Qy      1460 GATCCCGCACCCGAAAAACCTACCCGAAACCGAAAGCCA---AACTTTGGAACGC 1516
Db      428 rAlaPro-----ProThrThrSerArgProThrProAlaSerSerSerSerAl 444
Qy      1517 ATCGCGCGAAAAATCCTATCCCTCGTCG 1544
Db      444 aSerProProThrArgArgCys 453

RESULT 26
US-10-437-963-176617
; Sequence 176617, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176617
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(533)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74349C.1.pep
US-10-437-963-176617

```

Alignment Scores:	
Pred. No.:	3,728-11
Score:	292.50
Percent Similarity:	34.3%
Best Local Similarity:	26.8%
Query Match:	10.3%
DB:	4
<hr/>	
DB	QY
230	-----ProProLeuProProAspProAlaSerThrArgSerGlyAsn---245
1059	TGCAGGACGGCATAGACGTTACCGTCTGACCAACTCGCTACAGGCGACCGAGCTGTGCGG1118
246	-----ValAspProCysPro250
1119	CGGTCCATTCGGGCTACGTCAAATACCGAAAAACGGTGTCTCAAAGCGCGCATCAAAACTCT1178
	::: :::
	Length: 533
	Matches: 139
	Conservative: 39
	Mismatches: 179
	Indels: 163
	Gaps: 24


```
QY 1532 CTATCC 1537
Db 489 ValSer 490

RESULT 29
US-09-966-521-18
; Sequence 18, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-966-521-18

Alignment Scores:
Pred. No.: 9,27e-11 Length: 493
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 3 Gaps: 16

US-10-665-990A-13 (1-1561) x US-09-966-521-18 (1-493)

QY 110 GAAGAACGGACGAAGCGGTGATTTCAATCTTCCAAACCTGCTCCTCTGGAC----- 163
Db 104 AspAsnLysGlnValGlnLysHisAspLeuValArgMetLeuLeuMetAspGlnAsp 123
QY 164 -----AACATCTCGCAATCGGCACACCCCTCATATAACAACGGGCTATCCGAC 211
Db 124 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyrAsp 143
QY 212 ATCTACCTGCTCGACGACCCCGACGAAGCCCTTGCGCGCGCGCCCTTATCGAATCT 271
Db 144 GlnValLeuLysAsp-----IleLysAsn 151
QY 272 GCCGAACACAGCCTCGATTGCAATACATATTGGCGCAACGACATTTCCGGCAGGCTG 331
Db 152 AlaLysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 171
QY 332 CTGTTCAACCTCATGTACCTTCCCGCAGAACGGCGTGCGCTGCTGCTGTGGAC 391
Db 172 IleLeuHisAlaLeuGluGluLysLysGlnGlyLeuGluValLysIleLeuTyrAsp 191
QY 392 GACAAACAACCGCGGGGTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCCAATATC 451
Db 192 AspValGlySerLys-----AsnVal 198
QY 452 GAAGTGGCGCTTTCAACCCCTTGCTCTACGCAATGGCGGCGACTCGGC----- 502
Db 199 LysMetAlaAsnPheAspHisPhe-----LysSerLeuGlyGlyGluVal 213
QY 503 -----TACCTGACCGACTTCGCC-----CGCCTCAACCGCGCGACTGCAC 541
Db 214 GluAlaPhePheAlaSerLysLysLeuProLeuLeuAsnPheArgMetAsnAsnArgAsnHis 233
QY 542 AACAAATCTTTACCGCCGACAAACCGCGCCACCATACTCGCGGACGCAATATCGGCGAC 601
Db 234 ArgLysIleIleValIleAspGlyGlnLeuGluTyrValGlyPheAsnIleGlyAsp 253
```

RESULT 30

```
US-10-429-094-18
; Sequence 18, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
```

```
QY 602 GAATACTTCAAGTCGTGAGGACACACCGTTTTTCGCGACCTGGACATCCTCGCCACCGGC 661
Db 254 GluTyrLeuGlyLeuGlyLysLeuGlyTyrTrpArgAspThrHisLeuAlaGlnGly 273
QY 662 AGCGTCGTGCGCGAAGTATCGACGACTTCCACCGCTACTGGCAAGCAATTCGCCGCCAC 721
Db 274 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSer---GlnAlaHis 292
QY 722 AACGCCACGCCATCATC-----CGCAGCGGCAACATCGGCAAG 760
Db 293 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 311
QY 761 GGTCTTCAAGCACTCGGATACACGACGAACATCCACAGCGGCTCCTCGCTGCTACCGC 820
Db 311 ----- 311
QY 821 GAAACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGACGGGACGATCGACTGGCAG 880
Db 312 -----AsnSerProIleGlnIleAlaSerGlyProAlaSerAspTrpHis 327
QY 881 AGCGTCCAA-----ACCCGCTGATCAGCAGCACCCCTGCAAAAGGACTCGACCGC 931
Db 328 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 340
QY 932 GACGCCGCAACCGCCGATTGCGGGAGGCTGCAAGACGGCTCAACACGCCCGCAAAA 991
Db 341 -----LysLys 342
QY 992 AGCGTCTATCTGTTTACCCCTATTTCCTCCCTACAAAATCCGGCACAGACGCTGGCA 1051
Db 343 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 362
QY 1052 AAACCTGGTGACGACGCGATAGCTTACCTCTGCTGACCAACTGCTACAGCGGACCGAC 1111
Db 363 IleAlaAlaLysSerGlyValAspValHisMetIleProCys-----LysProAsp 380
QY 1112 GTTCCGCGCGTCCATTCCGGCTACGTCAAAATACCGAAAACCGCTCTCAAGCGGCGATC 1171
Db 381 HisProLeuValTyrTrpAlaThrPheSerAsnAlaSerAspLeuLeuSerSerGlyVal 400
QY 1172 AAATCTTACGAGCTGCAACCCCAACCATGCCGTCCCGCCACAAAAGACAAAGGCGCTGACC 1231
Db 401 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 409
QY 1232 GCGAGCTCCGTAACACGAGCTGATGCCAAAACCTTCATTGTGCGGCGCAAGCATCTTC 1291
Db 410 -----IleHisSerLysMetCysLeuIleAspAspGluIleValSer 423
QY 1292 ATCGGCTCATTCAACTCGACCCCGGTTCCGACGGCTCAATACCGAAATGGCGCTCGTC 1351
Db 424 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 443
QY 1352 ATCGAAAGCCCAAAATCGCAGAACAGATGGAGCGCACCCCTCGCGGATATCACACCCGAA 1411
Db 444 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 454
QY 1412 TAGCGCTACCGGTTTACCTCGCAAAACACACACCCCTGCAATGGCAGCATCCCGCCACC 1471
Db 455 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 468
QY 1472 CGAAAAACCTACCGAACGAACCGGACGCAAACTTTTGGAAACGCATCGCGCAAAAATC 1531
Db 469 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 488
QY 1532 CTATCC 1537
Db 489 ValSer 490
```


OTHER INFORMATION: expression in E. coli
US-09-966-521-84

Alignment Scores:

Pred. No.: 9,3e-11 Length: 502
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 3 Gaps: 16

US-10-665-990A-13 (1-1561) x US-09-966-521-84 (1-502)

```

QY 110 GAAGAACGACGGAAGCGTCAATTCAATCTCCAAACCTGCTCCTCTGGAC----- 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 AspAsnLysGlnValGlnLysHisAspLeuValArgMetLeuLeuMetAspGlnAsp 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 -----ACATCTCGAATCGGACACCCCTCTATAACAGGGCTATCCGAC 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyrAsp 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 ATCTACCTGCTCGACGACCCACGAGCCCTTGCCTGCGCCGCGCTTATCAATCT 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 145 GlnValLeuLysAsp-----IleLysAsn 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 272 GCCGAACACAGCCTCGATTTCATATCTACATTTGGCGCAACGACATTTCCGCGAGGCTG 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 AlalysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 CTGTTCAACCTCATGTACTTTCGCGCAGAACGCGCGTGCCTGACGCTCTGTGGAC 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 IleLeuHisAlaLeuGluGluLysLeuLysGlnGlyLeuGluValLysIleLeuTyrAsp 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 392 GACAAACACGCGCGGTGGACGATCTCTGTCTGCGCTCGACAGCCATCCCAATATC 451
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 AspValGlySerLys-----AsnVal 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 452 GAAGTGCCTGTTTCAACCCCTTCGTCTAGCGAAATGGCGCACTCGGC----- 502
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 LysMetAlaAsnPheAspHisPhe-----LysSerLeuGlyGlyVal 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 503 -----TACTGACCGACTTCCC-----CGCTCAACCGCGCATGCAC 541
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 GluAlaPhePheAlaSerLysLeuLeuLeuAsnPheArgMetAsnAsnArgAsnHis 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 542 AACAAATCTTTACCGCGCAACCGCGCCACCATCTCGCGGACGCAATATCGGCAC 601
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 ArgLysIleIleValIleAspGlyGlnLeuGlyTyrValGlyGlyPheAsnIleGlyAsp 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 602 GAATACTTCAAAGTCGGTGAGCACACCGTTTTCGCGCACTGGACATCTCGCCACCGGC 661
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 GluTyrLeuGlyLeuGlyLysLeuGlyTyrTrpArgAspThrHisLeuArgIleGlnGly 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 662 AGCTGTCGGCGAAGTATCCGACGACTTCGACCGCTACTGGCAAGCAATTCGCCAC 721
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 275 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSer---GlnAlaHis 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 722 AACCCACGCGCATCATC-----CCGACGGGCAACATCGGCAAG 760
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 761 GGTCTTCAAGCACTCGGATACAACGAGAAATCCAGACGCGCTCTCGCGTACCGC 820
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 312 ----- 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 821 GAAACCGTCGAACAGTCGCCCTCTACAAATAATACAGCGGACGCATCGATCGGCAG 880
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 -----AsnSerProIleGlnIleAlaAlaSerGlyProAlaSerAspTrpHis 328
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 881 AGCGTCCAA-----ACCGCTGATAGCAGACCCCTGCAAAAGAGACTCGCCGC 931
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 329 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 932 GACCGCGCAAAACCGCGGATTGTCGGGAGGTGCAAGACGCGCTCAAAACAGCCCGCAAAA 991
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 -----LysLys 343
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 992 AGCGTCTATCTGGTTTCAACCTATTTCCCTCCCTACAAAATCCGGCACAGACGCACTGGCA 1051
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 344 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1052 AAACCTGGTCAGGACGCGCATAGAGCTTACCGTCTCTGACCAACTCGCTACAGCGCACCGAC 1111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 IleAlaAlaLysSerGlyValAspValHisLeuMetIleProCys-----LysProAsp 381
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1112 GTTCCGCGCTCCATTCGGCTACGTCAATAATACCAAAACCGCTCTCAAAACCGCGCATC 1171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 382 HisProLeuValTyrTyrAlaThrPheSerAsnAlaSerAspLeuLeuSerSerGlyVal 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1172 AAACCTTACAGAGTGCACCAACCAACCATGCGCTCCCGCCCAAAAGAACAGGCTGACC 1231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1232 GGCAGCTCCGTAAACGCGCTGCATGCCAAACACCTTCATTGTGGAGCGGCAAAAGCATCTTC 1291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 -----IleHisSerLysMetCysLeuIleAspAspGluIleValSer 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1292 ATCGGCTCATTCACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGCGTCGTC 1351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1352 ATCGAAAGCCCCAAATCGCAGACAGATGGAGCGCACCCCTCGCCGATACACACCCGAA 1411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 445 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 455
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1412 TAGCCCTACCGCTTACCTCGACAAACACACCGCTCGAATGGCAGCATCCGCCACCC 1471
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1472 CGAAAAACCTACCGCAACCGAACCGAACCAACTTTTGGAAACGCATCGCGCAAAATC 1531
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 470 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1532 CTATCC 1537
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 ValSer 491
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 33

US-10-429-094-84
; Sequence 84, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774, US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned from
; OTHER INFORMATION: expression in E. coli
US-10-429-094-84

Alignment Scores:

Pred. No.: 9.3e-11 Length: 502

```
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 4 Gaps: 16

US-10-665-990A-13 (1-1561) x US-10-429-094-84 (1-502)

QY 110 GAAGAACGGACGGAAGCGTCATTTCATATCTTCAAACTGCTCTCTCTGGAC----- 163
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 105 AspAsnLysGlnValGlnLysHisAspLeuValArgMetLeuLeuMetAspGlnAsp 124
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 164 -----AACATCTGCAATCCGGCACACCCCTCAATAACACGGGTATCCGAC 211
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 125 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyrAsp 144
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 212 ATCTACCTCTCGACGACCCACCAAGGCCTTGGCGCGCGCCCTTATCAATCT 271
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 145 GlnValLeuLysAsp-----IleLysAsn 152
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 272 GCCGAACAGCCTCGATTGCAATACTACTATTGGCGCAACGACATTTCCGGCAGGTG 331
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 153 AlaLysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 172
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 332 CTGTTCAACCTCATGCTTCCCGCAGAACGCGGGTGGCGGTACGCTGCTGTGGAC 391
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 173 IleLeuHisAlaLeuGluLysLeuLysGlnGlyLeuGluValLysIleLeuTyrAsp 192
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 392 GACAACAACACGCGGGGTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATC 451
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 193 AspValGlySerLys-----AsnVal 199
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 452 GAAGTGGCGCTTCAACCCCTTGTCTTACGAAATGGCGGCGACTCGGC----- 502
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 200 LysMetAlaAsnPheAspHisPhe-----LysSerLeuGlyGlyGluVal 214
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 503 -----TACCTGACCGACTTCC-----CGCTCAACCGCGCGCATGCAC 541
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 215 GluAlaPhePheAlaSerLysLeuProLeuLeuAsnPheArgMetAsnAsnArgAsnHis 234
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 542 AACAAATCTTTACCGCCCAACACCGCCGACATCTCGCGGAGCAATATCGCGCAC 601
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 235 ArgLysIleIleValIleAspGlyGlnLeuGlyTyrValGlyPheAsnIleGlyAsp 254
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 602 GAATACTTCAAGTCGTGAGACACCGTTTTCGCGGACCTGGACATCTCCGCCACCGGC 661
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 255 GluTyrLeuGlyLeuGlyLysLeuGlyTyrTyrArgAspThrHisLeuArgIleGlnGly 274
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 662 AGCGTGTGCGGCAAGTATCGCAGACTTCCGCGCTACTGGCAGCCATTCGCCGCCAC 721
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 275 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTyrAsnSer---GlnAlaHis 293
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 722 AACGCCACGCGCATCATC-----CGCAGCGGCAACATCGGCAAG 760
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 294 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 312
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 761 GGTCTTCAAGCACTCGGATACACGACGAAACATCCAGACACGCGCTCTCGCTACCGC 820
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 312 ----- 312
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 821 GAAACCGTCGACAGTCCGCTTACCAAAAATACAGCGGACGCGCATCGACTGGCAG 880
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 313 -----AsnSerProIleGlnIleAlaAlaSerGlyProAlaSerAspTyrHis 328
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 881 AGCGTCCAA-----ACCGCGCTGTATCGACGACACCCCTGCAAAAGGACTCGACCGC 931
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 329 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 341
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 932 GACCGCGCAACCGCGCGATTGCGGGAGGCTGCAAGCGCGTCAACAGCGCCGCAAAA 991
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 342 -----LysLys 343
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 992 AGCGTCTATCTGTTTCAACCTATTTCGTTCCTTACAAAATCCGGCACAGACGCACTGGCA 1051
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
```

```
Db 344 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 363
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1052 AAATGTTGTCAGGACGCGCATAGCTTACCTGCTCTGCAACATCTGCTACAGCGCAGCCGAC 1111
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 364 IleAlaAlaLysSerGlyValAspValHisLeuMetIleProCys-----LysProAsp 381
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1112 GTTGGCGCGTCCATTTCGGGTACGTCAAATACCGAAAAACCGTCTCAAGCCGCGCATC 1171
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 382 HisProLeuValTyrTyrPheSerAsnAlaSerAspLeuLeuSerGlyVal 401
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1172 AAATCTACGAGTCAACCAACCATCGCTGCCCGCCCAAAAAAGCAAAAGCGCTGACC 1231
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 402 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 410
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1232 GGCAGCTCCGTAACCGAGCTGATGCCAAAACCTTCATTGTGGAGCGGCAACGCAATCTTC 1291
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 411 -----IleHisSerLysMetCysLeuIleAspGluIleValSer 424
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1292 ATCGGCTCATTCACCTGACCCCGTTCGGCAGCGGCTCAATACCGAAATGGCGCTGCTC 1351
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 425 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 444
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1352 ATCGAAAGCCCAAAATCGCAGACAGATGGAGCGCACCTCGCCGATACACACCCGAA 1411
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 445 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 455
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1412 TACGCTACCGGTTTACCTCGACAAACACAAACCGCTGCAATGCGCATGCCGCGCAC 1471
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 456 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 469
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1472 CGAAAAACCTACCCGAGCAAGCCGCAAGCAACCTTTCGAAACGCGATCGCCGCAAAATC 1531
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 470 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 489
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1532 CTATCC 1537
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 490 ValSer 491
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

RESULT 34
US-10-084-846A-3
; Sequence 3, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLNWEIG, AGNES
; APPLICANT: TREFFER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin Ver. 3.2
; SEQ ID NO 3
; LENGTH: 19695
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-3

Alignment Scores:
Pred. No.: 2,92e-10 Length: 19695
Score: 283.00 Matches: 175
Percent Similarity: 33.4% Conservative: 60
Best Local Similarity: 24.9% Mismatches: 226
Query Match: 9.9% Indels: 242
```

DB: 4 Gaps: 36
US-10-665-990A-13 (1-1561) x US-10-084-846A-3 (1-19695)
QY 80 TGTTCATGTTCTTCA---TGGTTG----- 100
DB 5065 CysSerCysSerSerAlaTrpSerValThrProThrMetLysAlaArgAlaAlaPro 5084
QY 101 ---CCCCACTGGAAGAGCGGAGGAAAGCCGTCATTTCATTAATCTTCCAAACCTGCTCCTC 157
DB 5085 GlyProPro-TripGluSerSerAlaSerAlaCysArgSer-----ProSerAlaValLe 5102
QY 158 CTGGACAACATCTCTCAATCGGCACACCCCTCATATAACAACGGGC-----TATCC 208
DB 5102 uTrpAspThrArgCysThrProAlaThrProArgThrAlaSerAlaSerCysArgSerPr 5122
QY 209 GACATCTACCTGCTCGACGACCCCAAGAGCCCTTGGCGCGCGCGCGCTTATCGAA 268
DB 5122 oSerSerTrpAlaSer-----ProAlaProProSerProCysTrpArgGl 5137
QY 269 TCTGCCGAACACAGCCTCGATTGCAATCTACTACATTTTGGCGCAACGACA----- 317
DB 5137 yCysTrpAsnAlaGly-----CysSerSerProGlyArgAlaProThrProSerTrpGl 5155
QY 318 -----TTTCGGCAGGCTGCTGTTCACCTCATGTATCTTCCGCGCAGAA 361
DB 5155 yProProProProlleSerProAlaGlyCysCysSerArgValSerSerProCysCysAl 5175
QY 362 CGCGCGTGGCGTACGCTGCTGTGGACGACA-----ACAACACGCGGGG 409
DB 5175 aAlaValArgProTrpAlaCysCysAlaProSerArgSerProArgSerSerAlaSe 5195
QY 410 TTGGACGATCTCTGCTCGCCC-----TCGACAGCCAT 442
DB 5195 rTrpTrpCysSerSerAlaProProCysAspAlaSerTrpSerGlyArgSerAlaGlyCy 5215
QY 443 CCCAATATCGAAGTGGCGCTGTTCAACCCCTTTCGCTTACGCCAAATGGCGCGCACTCGGC 502
DB 5215 sProThrAlaAlaCysSerHisSerSerProAlaSerCysCys----- 5229
QY 503 TACCTGACCGACTTCCCGCGCTCAACCGCGCATGCAACACAAAT----- 548
DB 5230 -----ProArgSerProProSerSerAlaSerIleSerSerSerAlaProSerCysSe 5247
QY 549 -----CCTTTACGCGCG-----AC 562
DB 5247 rAlaProSerAlaArgArgProArgArgProTrpArgThrProAlaAlaAsnCysArgAs 5267
QY 563 AACCGCGCCACATAC----- 578
DB 5267 nSerSerProProTrpCysCysArgSerSerProArgSerAlaArgProThrCysCy 5287
QY 578 ----- 578
DB 5287 sAsnSerAlaGlyAlaAlaProCysGlyCysGlyProAlaSerProCysTrpTrpPr 5307
QY 579 -----TCGGC 583
DB 5307 oSerSerAlaAsnTrpArgAlaAlaArgArgGlnArgArgCysArgTrpSerAlaSerTh 5327
QY 584 GGACGCAATATCGCGCAGATACT-----TC 610
DB 5327 rProCysAlaSerAlaCysSerThrAlaAlaAspProAsnTrpSerSerProSerGlySe 5347
QY 611 AAAGTCGGTAGGACACCG---TTTTCCCGACCTGGACATCTTCGCCACCGGCA-----GC 664
DB 5347 rAsnSerAlaSerHisProArgCysSerProCysTrpSerSerSerProCysAlaProPr 5367
QY 665 GTCCGCGCGCAAGTATCGACGACTTCGACCGCTACTGGGCAAGCC---ATTCCGCGCCAC 721
DB 5367 oSerProArgArgCysSerThrCysSerThrAlaProArgArgAlaProArgProProVa 5387
QY 722 AACGCCACGC-----GCATCATCGCAGCGGCAACATCGGCA----- 758

DB 5387 lGlnLysArgValProSerPheAlaAspAspGlyAlaAlaAsnSerAlaArgGluMetPr 5407
QY 759 ---AGGTCTTCAAGCACTCGATACAAACGACGAAACATCCAGACACGCGCTCCTCGCC 814
DB 5407 oGluArgValValIleProSerArgIleProThrLysArgLeuSerSerLysGluCysAr 5427
QY 815 TACCGCGAAACCGTCGACACGTCGCCCC----- 842
DB 5427 gSerValArgAlaGlySerSerLysProGluAsnArgGluValGlyTrpCysArgProTh 5447
QY 843 -----TCTACAAAAAATACAGACGGGACGACGCTCGACTGGC----- 878
DB 5447 rCysThrSerValLeuSerAlaSerSerTyArgArgValAlaSerAsnGlyProTrpAs 5467
QY 879 -----AGAGGTCCTCAACCGCTGTATCAGCGACACCCCTTGCAAAAGGACTC 925
DB 5467 nAlaAlaSerIleArgProLysArgProArgHisThrSerSerAlaGlySerProProAl 5487
QY 926 GACCGCGACCGCCCAACCGCCGATTCGCGGAGGCTGCAAGACGCGCTCAACAGCGCC 985
DB 5487 aThrSer-ProLeuArgLysTrpProCysGlyArgArgSerArgArgGlyGlyAlaA 5507
QY 986 GAAAAAAGCGTCTATCT-----GGTTTCACCC 1012
DB 5507 spArgArgArgAsnSerIleCysCysCysMetArgAlaProGlyThrArgAlaArgThrA 5527
QY 1013 TATTTCTGCTCTACAAAATCGGCACAGACGACATCGGCAAACTGGTGACAGCGGCATA 1072
DB 5527 laGlyArgPro-----ArgHisThrCysSerGlyThrTrpAlaGlyThrCysS 5544
QY 1073 GACGTTACCGTCTGACCACTCGCTACAGC-----GACC 1108
DB 5544 erProTrpArgSerGlyArgAlaAlaThrAlaCysSerAlaArgTrpAsnSerProSerA 5564
QY 1109 GACGTTCCCGCGTCCATTCGG-----CTACGTCATAATACCGAAACCGCTGTC 1159
DB 5564 laThrCys---ArgProIleArgThrAlaProAlaArgSerSerProProThrThrThrg 5583
QY 1160 AAAGCCCG-----CATCAACTCTACGA----- 1182
DB 5583 luArgArgSerThrGlyGlyGlyTrpAlaProAlaSerSerAlaAlaMetArgSerArgp 5603
QY 1183 -----GCTGCAACCCCAACCATGCGTCCCGCCCAACAAAGACAAAGCCCTGACCGCAGC 1237
DB 5603 roGlyAlaAlaSerGlyProAlaSerProGlySerAlaArgTrp---ProProArgGlyC 5622
QY 1238 TCGTAAACGCTGTCATGCAAAACCTT-----CATTTGGACGGCAACCGCATCTTC 1291
DB 5622 ysArgArgSerSerArgCysThrAlaValThrSerArgCysSerArgArg--AlaSerPr 5641
QY 1292 ATCGGCT---CATTCAACCTCGACCCCGCTTCGCGACGGCTCAATACGGAATGGGGTC 1348
DB 5641 oGlyAlaValProProThrSerAlaProValSerAlaSerSerSerProProAlaAlaPr 5661
QY 1349 GTCATCG---AAAGCCCCAAAATCGCAGAACAGATGGAGCGACCCCTCG---CCGATACC 1402
DB 5661 oProSerLeuArgTrpProArgPheArgThrIleTrpThrProSerAlaProTrpPr 5681
QY 1403 ACACCCG-----AATACGCTTACCGGTTTACCCCTCGACAAACACACCGCC-- 1448
DB 5681 oGlyProAlaSerGlyTrpArgThrSerProAlaSerProSerThrIleThrArgSerGl 5701
QY 1449 -----TGCAATGGCAGATCCCGCCACCCCGCAAAACCTACCCG 1486
DB 5701 yArgAlaValLeuAsnGlyGlyLeuGlyGlyAlaAlaValProValHisValGlyIleAr 5721
QY 1487 AACGAAACCGGAGCAAACTTTTGGAAACGATCGCGCAAAATCTCTATCCCTGCTGCC 1546
DB 5721 gProGlyAspArgProLeuAlaGlyProAlaThr-----IleCysCysPr 5736
QY 1547 A 1547

Db 5736 o 5736

RESULT 35

US-11-097-143-41094

; Sequence 41094, Application US/11097143

; Publication No. US2005020858A1

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; TITLE OF INVENTION: DROSOPHILA GENES.

; FILE REFERENCE: CL000728

; CURRENT APPLICATION NUMBER: US/11/097,143

; CURRENT FILING DATE: 2005-04-04

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 41094

; LENGTH: 2284

; TYPE: PR1

; ORGANISM: DROSOPHILA

US-11-097-143-41094

Alignment Scores:

Pred. No.:	2,79e-10	Length:	2284
Score:	281.00	Matches:	154
Percent Similarity:	34.1%	Conservative:	63
Best Local Similarity:	24.2%	Mismatches:	237
Query Match:	9.9%	Indels:	183
DB:	6	Gaps:	34

US-10-665-990A-13 (1-1561) x US-11-097-143-41094 (1-2284)

QY 80 TGTTCATGTTCTTCATGTTGTCCTCCCTGAGAACGACGCGGTCATTTCAT 139

Db 930 CysGlnProAlaThrTyrLeuProProSerThrLysArgThrProThr 946

QY 140 ACTTCCAAACCTGCTCTCTGACAAACATCTCGCAATCCGGCACACCCCTCATACAAAC 199

Db 947 AlaProLysPro-ArgCysTyrProGlySerLysAspProGlyCysProGlnIleThrTh 966

QY 200 GGGCTATCGACATCTACC-----TGC-----TCGACGACCCCC---AC 235

Db 966 rArgAlaProThrSerThrSerArgProArgCysTyrProGlySerThrAspProGluCy 986

QY 236 GAAGCCCTTGGCCCGCCCGCCCTTATCGAATCTGCGACACAGCCTCGATTGCAA 295

Db 986 shiProThrThrSerSerProAlaIleThrArgIleProValThrThrArgIleProLe 1006

QY 296 TACTACATTTGGCGAACACATTTCCGGCA-----GGCTGCTGTTCAACCTC 343

Db 1006 uThrThrAlaLysProArgCysTyrProGlySerGlnGluProGlyCys-----GlnPr 1024

QY 344 ATGTACCTTGGCGAACGCGCGGTGCGGTACGCGCTGC----- 383

Db 1024 oAlaThrArgProProIleThrThrSerLysProArgCysTyrProGlySerLeuAspPr 1044

QY 384 ----TGTTGACGACAACA-----ACAGCGCGGGTTGGAGCATCTCTCGTC 427

Db 1044 oGluCysGlnProSerThrTyrLeuProProThrProValArgThrThrValProThrIl 1064

QY 428 GCGCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTCGTCTCTAGCAAA 487

Db 1064 eProThrThrArgIleProValThr-----ThrSerLysProAsnCysTyrPro-- 1080

QY 488 TGGCGGCACCTCGGTACCTGACCGCACTTCCCGCCTCAACC-----GCCGATGC-- 539

Db 1081 -GlySerThrAspArgCysProLysGluProValThrThrProLysProArgCysTy 1100

QY 540 -----ACAACAATCCT-----TTACCGCGCAACCGCGCC 571

Db 1100 rProGlySerProAsnProGluCysGlnLysAlaThrTyrSerProProThrThrArgTh 1120

QY 572 ACCATACCTCGCGGACGCAATATCGGCGACGAATCTTCAAAGTCGGTGAGACACCGTT 631

Db 1120 rProValThrThrSerLysProAsnCysTyrProGlySerThrAspSerArgCysProGl 1140

QY 632 TTCGCGACCTGGACATCTCGCCA----- 656

Db 1140 nLysProProThrThrLeuLysProLysCysSerProGlySerSerAspProGluCysLe 1160

QY 656 ----- 656

Db 1160 uAsnCysTyrProGlySerProAspCysProLysValProThrThrLysLysSe 1180

QY 656 ----- 656

Db 1180 rGlyCysPheAspGlySerGlnAspProArgCysGlnProAlaThrTyrLeuProProSe 1200

QY 657 -----CCGCGAGGTGCTCGCGGAAGTATCGACAGACTTCGACCGCTACTGGGCA 706

Db 1200 rSerArgArgProProThrThrAlaProLysProArgCysHisProGlySerThrAspPr 1220

QY 707 AGCCATTCG--CCCACAACGCCGCGCATCTCCGCGCGCAACATCGGCAAGGT 763

Db 1220 oSerCysProGlnProThrGlnProIleThrThrArgLeuProIleThrThrGlnLysAl 1240

QY 764 CTTCAAGCACCTCGGATACACAGCAGCAACATCCAGACACGCGCTCTCGCTACCGCAA 823

Db 1240 aArgCysTyrProGlySerThrAla-----ProGluCysGlnProAlaThrThrProTh 1258

QY 824 ACCGTCGAACAGTCGCCCC----- 842

Db 1258 rProThrSerThrArgProIleThrThrIleThrThrSerLysProAsnCysTyrProGl 1278

QY 843 -----TCTACCAAAAAATACAGACGGGA 865

Db 1278 ySerThrAspArgCysProGlnIleProValThrThrThrLysProArgCysTyrPr 1298

QY 866 GCGATCGACTGGCAGAGGTCCAAACCGCGCTGATACGACACCCCTGCAAAAGACTC 925

Db 1298 oGlySerThrAspProGlnCysGlnPro-----AlaThrTyrLeuProProThrTh 1315

QY 926 GACCGGACCCCGCAACCCCGCATTCGCGGA-----GGCTGCAAG----- 968

Db 1315 rValThrThrValArgProArgCysTyrProGlySerAsnAspProGlyCysGlnProPr 1335

QY 969 -ACGGCTCAAAACAGCCGCAAAAAAGCGCTATCTGTTTCACCTATTTCGTCCTCA-- 1025

Db 1335 oGlnThrThrLysThrProIleThrThrSer-----LysProValCysTyrLeuGl 1352

QY 1026 -----CAAAATCCGCGACAGCAGCAGCTCGGCAAACTGTGTGAG 1063

Db 1352 ySerThrAspSerArgCysArgGlnAsnProSerThrThrGln---LysProLysCysSe 1371

QY 1064 GACGCGATAGAGTTACCGTCC---TGACCAACTCGCTACAGCGGACCG--ACGTTGCC 1117

Db 1371 rValGlySerThrAspProGluCysGlnProAlaThrTyrLeuProProThrThrValSe 1391

QY 1118 GCGGTCCATTCCGGCTAGCTCAAAATACGCAAAACCGCTCTCAAAAGCGGATCAAACTC 1177

```
Db 1391 rProThrIlePro-----ThrThrLysProArgCys---TyrProGlySerThrAs 1407
|||:::||||| |||::: ||||| ||| |||
Qy 1178 TAGAGCTGCAACCAACCATCGCGTCCCGCCA-----CA 1213
||||| ||| ||| ||| ||| ||| ||| |||
Db 1407 pThrSerCysGlnProLysThrTyrSerProProThrSerArgProLeuIleThrThrSe 1427
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1214 AAAGCAAAAGCGCTGACCGGAGCTCCGTAACACCGCTGCATGCCAAAACCTTCATTGTG 1273
|||::: ||| ||| ||| ||| ||| ||| |||
Db 1427 rLysProArgCysTyrProGlySerProAspProGluCysGlnProAlaThrTyrLeu-- 1446
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1274 GACGGCAACGATCTTCATCGGCTCATCACTCGACCCCGCTCGCAGCGCTCAAT 1333
|||::: ||| ||| ||| ||| ||| ||| |||
Db 1447 -----ProProThrThrArgArgThrThrIleProVal-----ThrLy 1459
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1334 ACCGAATGGCGTGCATCGAAAGCCCAAAATCGCAG--AACAGATGGAGCGCAC 1390
|||::: ||| ||| ||| ||| ||| ||| |||
Db 1459 sProArgCysTyrProGlySerSerAspProGluCysGlnProGluThrTyrSerProPr 1479
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1391 CTCGCCGATACCAACACCGAATAG-----CCTACCGCTTACCTCGACAAACACAAC 1444
|||::: ||| ||| ||| ||| ||| ||| |||
Db 1479 oThrArgProProValThrThrSerLysProAsnCysTyrProGlySerThrAspSe 1499
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1445 CGCTGCAATGGCAGATCCCGCACCCGAAACCTTACCGAAGCAACCGAAGCCAAA 1504
|||::: ||| ||| ||| ||| ||| ||| |||
Db 1499 rArgCys-----ProGlnLysProProThrThrLeuGlnProLysCysTy 1514
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1505 CTTTGGAAACGCATCGCGCAA-----AAATCTATCCCTGCTGCCCA 1547
|||::: ||| ||| ||| ||| ||| ||| |||
Db 1514 rProGlySerSerAspProGluCysLeuAsnCysTyrProGlySerPro 1530
|||::: ||| ||| ||| ||| ||| ||| |||

RESULT 36
US-10-684-422-44
; Sequence 44. Application US/10684422
; Publication No. US20040229233A1
; GENERAL INFORMATION:
; APPLICANT: ABURATANI, Hiroyuki
; APPLICANT: YAMAMOTO, Shogo
; TITLE OF INVENTION: Human housekeeping genes and human tissue-specific genes
; FILE REFERENCE: 113991
; CURRENT APPLICATION NUMBER: US/10/684,422
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/418,614
; PRIOR FILING DATE: 2002-10-16
; NUMBER OF SEQ ID NOS: 332
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-684-422-44

Alignment Scores:
Pred. No.: 2,55e-10 Length: 820
Score: 280.50 Matches: 134
Percent Similarity: 37.8% Conservative: 59
Best Local Similarity: 26.3% Mismatches: 204
Query Match: 9.8% Indels: 113
DB: 5 Gaps: 23

US-10-665-990A-13 (1-1561) x US-10-684-422-44 (1-820)

Qy 217 CTGCTCGAGACCCCAAGAGCCCTTGGCGCCGCGCCCTTATCGAATCTGCGCA 276
||||| ||| ||| ||| ||| ||| ||| |||
Db 172 ProArgArgLysSerArgSerProSerProArgArgSerProValArgArg 191
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 277 ACACGCTCGATTGCACTACTACATTGGCGCAACGACATTTCCGGCAGGCTGCTGTT 336
|||::: ||| ||| ||| ||| ||| ||| |||
Db 192 GluArgLysArg-----SerHisSerArgSer----- 200
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 337 CAACCTCATGTACTTCGCGAGAACCGCG--CGTGGCGGTACGCTGCTGTGGACGA 393
|||::: ||| ||| ||| ||| ||| ||| |||
Db 201 -----ProArgHisArgThrLysSerArgSerProAlaProGluLys 216
|||::: ||| ||| ||| ||| ||| ||| |||
```

```
Qy 394 CAACAAACACGCGGGTTGGACGATCTCTCTGCTCGCCCTCGACAGCCATCCCAATATCGA 453
|||::: ||| ||| ||| ||| ||| ||| |||
Db 217 LysGlu-----LysThrProGluLeuProSerValLysValLys 231
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 454 AGTGGCGCTGTTCACCCCTTCGTCTACG---CAAAATGGCGCGCCTCGCTACCTCGAC 510
|||::: ||| ||| ||| ||| ||| ||| |||
Db 232 GluProSerValGlnGluAlaThrSerThrSerAspIleLeuLysValProLysProGlu 251
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 511 CGATTTCCTCCCTCAACCGCCGATGCAACAATCTCTTACCCTCCGCAACCGCGC 570
|||::: ||| ||| ||| ||| ||| ||| |||
Db 252 ProIleProGluProLysGluProSerProLysLysAsnSerLysLysGluGlnGluLys 271
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 571 CACCATCTCTGG---CGAGCGCAATATCGCGCAGCAATATCTTCAAGTCGCTGAGGACAC 627
|||::: ||| ||| ||| ||| ||| ||| |||
Db 272 GluLysThrArgProArgSerArgSerLys-----SerArgSerArgThr 288
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 628 CGTTTTCGCGACCTGGACATCTCTCGCCACCGCAGCGT----- 666
|||::: ||| ||| ||| ||| ||| ||| |||
Db 289 ArgSerArgSerProSerHisThrArgProArgArgHisArgSerAspLysMetTyr 308
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 667 -----CGTGGCGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCCATTCGCGCA 720
|||::: ||| ||| ||| ||| ||| ||| |||
Db 309 SerProArgArgProSerProArgArgProSerProArgArgThrProPro 328
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 721 CAACGCCACGCGCATCTCCGACGCGCAACATCGCAAGGCTCTTCAAGCACTCGGATA 780
|||::: ||| ||| ||| ||| ||| ||| |||
Db 329 ArgArgMetPro-----ProProArgHisArgSerArgSerProValArg 346
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 781 CAACGACGA-----AACATCCAGACA 801
|||::: ||| ||| ||| ||| ||| ||| |||
Db 347 ArgArgArgSerSerAlaSerLeuSerGlySerSerSerSerSerSerSerArgSer 366
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 802 CGCGCTCTCGGTACCGCGAAACCGTGCAGACGTCCGCCCTCTTACCAAAAATACAGAC 861
|||::: ||| ||| ||| ||| ||| ||| |||
Db 367 ArgSerProLysLysProProLysArgThrSerSerProProArgLys----- 383
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 862 GGGACGATCGACTGGCAGAGCGTCCAAACCCGCTCATCAGCAGAC----- 909
|||::: ||| ||| ||| ||| ||| ||| |||
Db 384 ---ThrArgLeuSerProSerAlaSerProArgArgArgHisArgProSerPro 402
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 910 -----CCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATTGCGCG 957
|||::: ||| ||| ||| ||| ||| ||| |||
Db 403 ProAlaThrProProLysThrArgAspSerProThrProGlnGlnSerAsnArgThrArg 422
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 958 GAGCTGCAAGACGCGCTCAACAGCCCGA----- 987
|||::: ||| ||| ||| ||| ||| ||| |||
Db 423 LysSerArgValSerValSerProGlyArgThrSerGlyLysValThrLysHisLysGly 442
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 988 ---AAAAAGCGTCTATCTGTTTCCACCTATTTGCTCCCTACAAATC----- 1032
|||::: ||| ||| ||| ||| ||| ||| |||
Db 443 ThrGluLysArgGluSerProSerProAlaProLysProArgLysValGluLeuSerGlu 462
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1033 ---CGGCACAGACGCTGGCAAACTGGTGCAGGA-----CGCATAGACGT 1077
|||::: ||| ||| ||| ||| ||| ||| |||
Db 463 SerGluGluAspLysGlyGlyLysMetAlaAlaAspSerValGlnGlnArgArgGln 482
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1078 TACGCT-----CCTGACCAACTCGCTACAGCGGACCGAGCTTGC 1116
|||::: ||| ||| ||| ||| ||| ||| |||
Db 483 TyrArgArgGlnAsnGlnSerSerSerAspSerGlySerSerSerSerSerGluAsp 502
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1117 CGCGCTCATTCGCGCTACGT---CAAAATACCGAAACCGCTGCTCAAGCCGCGCATCAA 1173
|||::: ||| ||| ||| ||| ||| ||| |||
Db 503 GluArgProLysArgSerHisValLysAsnGlyGluValGlyArgArgArgHisSer 522
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1174 ACTCTAGAGCTCAACCCCAACCATGCGCTCCCGCCACCAAAAGACAAAGCCCTGACCGG 1233
|||::: ||| ||| ||| ||| ||| ||| |||
Db 523 ProSerArgSerAlaSerProSerProArgLysArgGlnLysGluThrSerProArgGly 542
|||::: ||| ||| ||| ||| ||| ||| |||
Qy 1234 CAGCTCGTTAACAGCTGTCATGCCAAAACCTTCATTGTGGAGCGGCAACCGCATCTTCAT 1293
|||::: ||| ||| ||| ||| ||| ||| |||
Db 543 ArgArgArgArgSerProSer-----ProProThrArgArgArg 556
|||::: ||| ||| ||| ||| ||| ||| |||
```



```
Db 344 AlaAla-----SerArgPro---ProArgThrArgValProProAlaArgAla 359
QY 1353 TCGAAGCCCAAAATCGAGAACAGATGGAGCGCACCTCGCGATACACACCCGAAT 1412
Db 360 SerProSerArgGlyArgSerArgTrpArgProAlaTrpSerGlyAlaProGly 379
QY 1413 ACGCTACCGGCTTACCTCGACAAACACACCGCTGCAATGGCAGCATCCCGCACCC 1472
Db 380 CysProProCysSerProGlySerIleAlaThrSerPheSerCysThrProProPro 399
QY 1473 GAAAAACCTACCGGAACGAAACCCGAAG 1499
Db 400 ProProArgArgArgGlyPheProArg 408

RESULT 38
US-10-437-963-134726
; Sequence 134726, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134726
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36471C.1.pep
US-10-437-963-134726

Alignment Scores:
Pred. No.: 3,91e-10 Length: 668
Score: 277.50 Matches: 145
Percent Similarity: 33.9% Conservative: 57
Best Local Similarity: 24.4% Mismatches: 222
Query Match: 9.7% Indels: 171
DB: 4 Gaps: 30

US-10-665-990A-13 (1-1561) x US-10-437-963-134726 (1-668)
QY 33 CCCCGATGAAGAACGAGCGAGCTCATTTCCCTTTATGCC-----TCC 74
Db 43 ProThrGlyGlnHisHisSerSerProSerArgSerTyrlleThrGlnSerLeuPro 62
QY 75 TTCCTCTGATGTCTTCATGGTTGCCCCACTGGAAGACGACGAGGAAAGCGTCATT 134
Db 63 TyrleulleuLeuLeuLeuGlyTyrlValAsnTrpSerAsnGlyGlyGly----- 80
QY 135 TCAATACTTCAAAACCTGTCTCTCTGGACACATCTCTGCAATCCGGCACACCCCTCATA 194
Db 80 ----- 80
QY 195 ACAACGGGCTAT-----CCGACATCTACCTGCTCGAGGACCCGCCAGAGCCC----- 242
Db 81 -----GlyTyAsnProProSerProSerIleGlyThrSerProThrThrProGlyGly 98
QY 243 -----TTGCCGCGCGCGCGCCCTTATCGAATCTG 272
Db 99 GlyGlyGlyTyrlThrProThrProSerAspThrProProSerProSerAspThrSer 118
QY 273 CCGAACACAGCCTCGATTGTCATTTGCGCGCAACGACATTTCCGGCAGGCTGC 332
```

```
119 ProSerThr-----ProGlyGlyGly 125
QY 333 TGTTCACCTCA-----TGTACCTTGCCTGGAGAACGCGCGTGGCGTACGCCGTGC 383
Db 126 CysSerSerProThrProCysAspAlaProProSerProSerProSerAspThrSer--- 144
QY 384 TGTTCGACGACAAACACGCGCGGTGGAGCATCTCTCTCGCCTCGACGCCATC 443
Db 145 -----ProThrThrProGlyGlyGlyTyrlSerProThrProSerAspThrPro 162
QY 444 CCAATATCGAAGTCGCGCTGTTC----- 467
Db 163 ProSerProSerSerAspThrSerProThrThrProGlyGlyGlyGlyTyrlThrPro 182
QY 468 ACCCTTCGCTCTAGCAAAATGGCGGACCTGGCTACCTGACGACTTCCCGCGCTCA 527
Db 183 ThrProSerAspAlaProProSerProSerAspThrSerProThrThrProGlyGly 202
QY 528 ACCGCGCATGACACAAACAAATCCT-----TTACCGCGGACA-----ACGCG 569
Db 203 GlyGlyGlyTyrlThrProThrProSerAspAlaProProSerProSerSerAspThrSer 222
QY 570 CCA-----CCATACCTCGCGGACGCAATATCGCGC 599
Db 223 ProThrThrProGlyGlyGlyTyrlThrProThrProSerAspThrProProSer 242
QY 600 ACGAATACTCAAAAGTCGGTAGACACCG-----TTTCGCCGACCTGG 644
Db 243 ProSerSerGlySerSerProThrThrProGlyGlyGlyGlyTyrlThrProPro 262
QY 645 ACATCTCGCACCGGAGCGTCTCGCGAAGTATCGACGACTTCGACCGCTACTGGG 704
Db 263 SerAspThrProProSerProSerSerGlySerSerArgThrThrProGlyGlyCysSer 282
QY 705 CAAGCCATTCCGCCCAACACGCCGCGCATCTCCGCGGCAACATCGGCAAGGTC 764
Db 283 ThrProThrProCysGlyThrProProAlaProSerSerGlyThr----- 297
QY 765 TTCAGCAGCTCGGATACAAG-----ACGAAACATCCAGACACGCGCTCTCGCGCT 815
Db 298 -----SerProThrThrProGlyGlySerTyrlProProThrProSerIleGly 314
QY 816 ACCGCGAAACCGTCGAACAGTCGCCCTCTACCAAAATAACAGACGGGACGATCGACT 875
Db 315 AspValProProSer-----ProSerSerAspThrSerProThrThrProGlyGly 331
QY 876 GGCAGAGCGTCCAAACCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGGACC 935
Db 332 GlySerProSerThrProCys-----AspThrPro-----ProSerProSerSer 346
QY 936 GCGCAACCGCGATTCGCGGAGGCTGCAAGCGGCTCAACACGCCGCAAAACGCG 995
Db 347 GlyThrSerProThrThrProGlyGly-----GlyTyrlProProThrPro 362
QY 996 TCTATCTGGTTTCACCTATTTCG-----TCCCTACAAATCCGGCACAG 1040
Db 363 SerValGlyAspValProProSerProAlaSerGlyThrSerProSerThrProGlySer 382
QY 1041 ACGCACTGGCAAAACTGGTCAGGAGGATAGACGTTACCGTCTCTGACCACTCGCTAC 1100
Db 383 GlyGlyTyrlSerPro-----SerThrProCysSerAlaProProSerProSerGly 400
QY 1101 AGCGGACCGAGTTGCG-----CCGTCCATTCCGGCTACGTCAATAATACC 1145
Db 401 ThrSerProThrThrProGlyGlyTyrlSerProSerThrPro----- 415
QY 1146 GAAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCAACCAATGCGCTCC 1205
Db 416 -----CysAsnAlaProProSerProSerSerAspThrSerProThrThrProGly 432
QY 1206 -----CCGCCACAAAGACA-----AAG 1223
```

Db 433 GlyGlyAsnTyrProProAlaProThrIleGlyAsnValProProSerProSerSerSerGly 452
QY 1224 GCCTGACGGCAGCTCCGTAACACAGCTGCATGCCAAACACCTTCAATTGTGGACGGCAAAAC 1283
Db 453 ThrSerProSerThrProGlyGlyCysSerSerProThrProCysAspAlaPro 472
QY 1284 GCATCTCATCGGCTCATTAACCTCGACCCCGCTTCGCGACGGGTCA----- 1331
Db 473 ProSerProSerSer---AspThrSerProThrProGlyGlyGlyTyrTyrProPro 491
QY 1332 ATACCGAAATGGCGCTCTCATCGAAAGCCCAAAATCGCAGAACAGATGGAGCGCACCC 1391
Db 492 ThrProSerIleGlyThrSerProSerThrProGlyThrGlyGlyGlyTyrTyrProPro 511
QY 1392 TCGCGG-----ATACACACACCGAATAGCGCTTACCGCGTTACCTCTCGACA 1436
Db 512 SerProSerThrGlyGlyTyrThrProThrProAspValProIleSerThrProSer--- 530
QY 1437 AACACACCGCTCGAATGGCAGCATCCGCCACCCGAAACCTTACCCGAAACGACCGG 1496
Db 531 -----SerProTyrSerProLeuValProThrPro-----ProSerSerThrThrPro 546
QY 1497 AAGCCAAACTTTGGAAACGCATCGCGCGCAAAATCCTATCCCTGC 1541
Db 547 MetProPheAspProAsnThrAlaPro-----PheProCys 558

RESULT 39

US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLWEG, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8
; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

Alignment Scores:
Pred. No.: 6,83e-10 Length: 19608
Score: 277.50 Matches: 140
Percent Similarity: 38.9% Conservative: 67
Best Local Similarity: 26.3% Mismatches: 222
Query Match: 9.7% Indels: 103
DB: 4 Gaps: 23

US-10-665-990A-13 (1-1561) x US-10-084-846A-8 (1-19608)

QY 104 CCATCGGAAGACGGACGAAACCGCTCATTTCAATA----- 140
Db 5104 ProLeuGluArgAspTyrThrArgArg-SerIleSerGlnLeuProValProAspVa 5123
QY 141 -----CTTCCAAACCTCTCTCTGGACAACATCTCGCAATCCGGCACACCCCTCAT 193
Db 5123 lHisPheTyrProGlnCysAlaAlaThrAsnLeuAsnTyrGlnTyrGlyAlaProCys-- 5142
QY 194 AACACGGGCTATCCGACATCTACTGCTCGAGACCCCGACGAGCCCTTCTCGCGCCGC 253

Db 5143 -----ProLysLysCysSerSerAlaValGlyArgArgAspProArgAr 5157
QY 254 GCCGCCCTTATCGAATCTGCCGAACACAGCTCCGATTTGCAAT-----AC 298
Db 5157 gProProCysSerAlaArgProAlaCysProSerTyrTrpAsnGlyArgProSerArgAl 5177
QY 299 TACATTTGGCCAAACGACATTTCCGGCAGGCTGTGTTCAACCTCATGTACTTGCCTGCCGA 358
Db 5177 aThrThrSerAlaSerArgSerProProAlaAlaProSerSerThrSer-SerAlaA 5197
QY 359 GAAACGGCGTGGCGGTACGCTCTCTTGGACGACAAACACACGCG----- 405
Db 5197 rgtTArGArgSer---ThrProAlaValThrArgSerArgThrAlaCysCysAlaG 5216
QY 406 -----CGGTTGGACGATCTCTGCTCCCTCGACAGCCATCCCAATATCGAAGTG 457
Db 5216 lAlaAArgArgThrGlyArgSerThrGlyArg--ArgSerSerValProThrSer----- 5233
QY 458 CGCTGTTCAACCCCTTCTCTACGNAATGGCGGCACTCGGCTACTCTACCGACTTC 517
Db 5234 -----AlaProGlyArgSerThrAlaThr----- 5241
QY 518 CCCGCTCAACCGCGCATGCACAAACATCCTTACCGCGCAGCAACCGCGCCACCAT 577
Db 5242 -ThrSerThrCysSerArgThrArgSerSerArgAlaProThrProArgAlaSe 5261
QY 578 CTCGGCGGACGCAATATCGCGACGAAATACTTAAAGTCGGTGAGGACACCGTTTCGCC 637
Db 5261 rArgSerSerAlaCysCysSerThrAlaThrAlaArgLeuGlyArgSerGlyTyrThrPr 5281
QY 638 GACCTGGACATCTCGCCACCGGAGCGCTCTCGGGAAGTATCGCAGCATTCGACCGC 697
Db 5281 oThr-----ProValSerTyrSerAlaArgSerSerThrThrSer---Se 5295
QY 698 TACTGGGCAAGCATTTCGCCCAACCGCCAGCGCATCATCCGCGGCGGCAACATCGGC 757
Db 5295 rThrProArgAlaGlyProAlaSerSerProSerThrArgThrAlaAlaArgThrArgSe 5315
QY 758 AAGGTCTTCAAGCCTCGGATACAAACGACGAAACATCCAGACACGCGCTCTCGCTAC 817
Db 5315 rSerAlaThrSerArgSerGlyAlaThrGlyArg-----AlaAlaArgCysCysPr 5332
QY 818 CGCGAAACCGTCGAACAGTCGCCCTCTACCAAAAATACAGACGGGACGATCGACTGG 877
Db 5332 oValArgProGlnGlyAla-----SerThrSerSerAlaProArgThrAlaGlyThrGl 5350
QY 878 CAGAGCGTCCAAACCGCTCATGACGACACCCCTGCAAAAGGACTCGACCGCGCACGC 937
Db 5350 y-----SerArgCysAlaGlyThrGlyThrAlaSerValSerSerAlaThrArgThr-- 5367
QY 938 CGCAACCGCGATTGCCGGAGGCTGCAAGACGCGCTCAACACGCGCGCAAAACGCTC 997
Db 5368 -----GlySerTyrSerAlaAlaGlyAsnThrProThrTyrArgAr 5381
QY 998 TATCTGTTTCACTTATTTGTCCTACAAAATCCGCGACAGACGCTGCGCAA----- 1052
Db 5381 gCysTrp---ProArgTrpSerArgAsnProArgProCysAlaThrSerTyrArgThrAl 5400
QY 1053 -----AACTGGTCGAGCAGCGCATAGAGTTCCTCGCTCTGACCAAC 1093
Db 5400 aSerThrSerArgAlaTyrGlyTyrSerArgThrSer-----ProThrSerProTh 5417
QY 1094 TCGCTACAGGCGACCG---ACGTTGCCCGCTTCATTCGGGTACGTCAATATACGCAAAA 1150
Db 5417 rAlaSerThrAlaProAlaThrThrSerProAlaThrProProAlaSerSerThr----- 5435
QY 1151 CGCTGCTCAAGCGCGCATCA-----AACTCTACGAGCTCAACCCCAACCATGCC 1201
Db 5436 -ArgCysCysProProAlaSerThrTyrArgSerThrAlaAlaCysTyrProArgArgPr 5455
QY 1202 GTCCCGCGCA-----CAAAAGACAAAGCGCTG 1228

```

Db 5455 cYssSerProArgThrArgThrSerArgArgArgProGlyArgSerThrSerArgSe 5475
QY 1229 ACCGCAGCTCGTAACAGCCTGCATGCCAAACCTTCATTGTGGACGGCAACGCATC 1298
Db 5475 rThAlaThrProThrSerGlyCysSerProTrpCysProAlaCysThrSerAsnArgPr 5495
QY 1289 TTCATCGGCTCATCAACCTCGAACCCCGTTCCGACCGCTCAATPACCGAAATGGCGTC 1348
Db 5495 oAlaArgProProThrSerSerTrpProThrArgProArgSerAlaProArgProSerTh 5515
QY 1349 GTCATCGAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCG-----CGATACCA 1405
Db 5515 rSer-----ArgSerThrAlaArgTrpProSerProAsnTrpSerArgAspTr 5531
QY 1406 CCCGAATACGCTACCGGTTACCTCGACAAACACACCCGCTGCAATGGCACGATCCC 1465
Db 5531 pProThrSerThrProSerProAlaArgThrArgThrAlaProSerArgArgThPr 5551
QY 1466 GCCACCCGAAACCTACCCGAAACGAAACCCGAAACCTTTGGAAAC-----GCATC 1519
Db 5551 oArgProThrProCysArgArgThrThrProSerAlaSerCysSerArgArgProArgTh 5571
QY 1520 GCCGCAAAATCCTATCCTCGTGCCTCA 1547
Db 5571 rProValTrpArgArgProGlyCysPro 5580

```

RESULT 40

```

US-10-335-977-9083
; Sequence 9083, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 9083:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...206

```

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 9083:
US-10-335-977-9083

```

```

Alignment Scores:
Pred. No.: 5,12e-10 Length: 206
Score: 274.50 Matches: 64
Percent Similarity: 53.4% Conservative: 45
Best Local Similarity: 31.4% Mismatches: 84
Query Match: 9.6% Indels: 11
DB: 4 Gaps: 5

US-10-665-990A-13 (1-1561) x US-10-335-977-9083 (1-206)
QY 947 CCGATTCCGGGAGGCTGCAAGACGCTCAACAGCCGCAAAAGCGTCTATCTGTT 1006
Db 6 ProllelylleAlaPheGluLysAlaLeuLysAsnAlaLysGluSerValPheIleAla 25
QY 1007 TCACCTATTTCGTCCTTACAAAATCCGCGCACACAGCAGCTGGCAAACTGGTCAGGAC 1066
Db 26 SerSerTyrPheIleProGlyLysLysIleMetLysIlePheLysAsnGlnIleSerLys 45
QY 1067 GGCATAGAGTTACCGTCTGACCAACTCGCTACAGGCGACCGAGCTTCCGCGCTCCAT 1126
Db 46 GlyIleGluLeuAsnIleLeuThrAsnSerSerLeuSerThrAspAlaIleValValTyr 65
QY 1127 TCCGGCTACGTCAATACCGAAACCGCTGCTCAAGCCGCGCATCAAACTCTACGAGCTG 1186
Db 66 GlyAlaTrpGluArgTyrArgAsnLysLeuValArgMetGlyAlaAsnValTyrGluIle 85
QY 1187 CAACCCAACTGCGCTCCCGCCCAAAAGAGCGCTGACCGGAGCTCCGCTAAC 1246
Db 86 ArgAsnAspPhePheAsnArgGlnIleLysGlyArg-----PheSerThrLysHis 102
QY 1247 AGCTGATGCGCAAAACCTTCATTGTGCGAGCGGCAAAACGATCTTCATCGGCTCATTAAC 1306
Db 103 SerLeuHisGlyLysThrIleValPheAspAlaLeuThrLeuLeuGlySerPheAsn 122
QY 1307 CTCGACCCCGTTCCGCGACGCTCAATACCGAAATGGCGTCTCGTCAATCGAAAGCCCAA 1366
Db 123 IleAspProArgSerAlaTyrIleAsnThrGluSerAlaValLeuPheAspAsnProSer 142
QY 1367 ATCGCAGACAGATGGAGCGACCCCTCCGCGATACCAACACCCGGAATACGCTACCGGCT 1426
Db 143 PheAlaLysArgValArgLeuSerLeuLysAspHisAla---GlnGlnSerTrpHisLeu 161
QY 1427 ACCCTCGACAAACACACCGCTGCAATGGACGATCCCGCCACCCGAAACAACTAC--- 1483
Db 162 ValLeuTyrArgHis---ArgValIleTrpGlu-----AlaThrGluGluGlyIleLeu 178
QY 1484 -----CCGAACGAACCCGAGGCCAAACTTTTGAACAGCATCGCGCGCAAAATCCTA 1534
Db 179 IleHisGluLysAsnSerProAspThrSerPhePheLeuArgLeuIleLysGluTrpSer 198
QY 1535 TCCTGTGTCGCC 1546
Db 199 LysValLeuPro 202

```

```

Search completed: May 2, 2006, 06:06:03
Job time : 426 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:57:05 ; Search time 6.3 Seconds
(without alignments)
2253.681 Million cell updates/sec

Title: US-10-665-990A-13

Perfect score: 2852

Sequence: 1 caaatacaggaatgcgt.....tgccatcgagggtttatta 1561

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10665990/runat_01052006_112005_9866/app_query.fasta.1
-DB=Published_Applications_AA_New -QFWT=fasten -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCOR=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10665990.CG1.1.34 @runat_01052006_112005_9866 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:*

1:	/SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep.1.*
2:	/SID55/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3:	/SID55/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4:	/SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5:	/SID55/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
6:	/SID55/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
7:	/SID55/ptodata/2/pubpaa/US09_NEW_PUB.pep.1.*
8:	/SID55/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
9:	/SID55/ptodata/2/pubpaa/US10_NEW_PUB.pep.1.*
10:	/SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
11:	/SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep.1.*
12:	/SID55/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2681	94.0	525	9	US-10-467-657-5462
2	336.5	11.8	5179	11	US-11-108-172-1068
3	294.5	10.3	504	11	US-11-045-004-2518
4	287	10.1	482	11	US-11-045-004-2698
5	250.5	8.8	1970	9	US-10-821-234-1641

6	249	8.7	702	11	US-11-096-568A-14646	Sequence 14646, A
7	248	8.7	493	11	US-11-096-568A-11657	Sequence 11657, A
8	247.5	8.7	941	11	US-11-169-232-14	Sequence 14, Appl
9	247.5	8.7	1022	11	US-11-169-232-84	Sequence 84, Appl
10	247.5	8.7	1038	11	US-11-169-232-74	Sequence 74, Appl
11	247.5	8.7	1049	11	US-11-169-232-58	Sequence 58, Appl
12	247.5	8.7	1140	11	US-11-169-232-104	Sequence 104, Appl
13	247.5	8.7	1270	11	US-11-169-232-44	Sequence 44, Appl
14	247.5	8.7	1311	11	US-11-169-232-42	Sequence 42, Appl
15	247.5	8.7	1313	11	US-11-169-232-142	Sequence 142, Appl
16	247.5	8.7	1314	11	US-11-169-232-50	Sequence 50, Appl
17	247.5	8.7	1320	11	US-11-169-232-46	Sequence 46, Appl
18	247.5	8.7	1320	11	US-11-169-232-60	Sequence 60, Appl
19	247.5	8.7	1354	11	US-11-169-232-48	Sequence 48, Appl
20	247.5	8.7	1361	11	US-11-169-232-40	Sequence 40, Appl
21	247.5	8.7	1363	11	US-11-169-232-52	Sequence 52, Appl
22	247.5	8.7	1404	11	US-11-169-232-2	Sequence 2, Appl
23	247.5	8.7	1404	11	US-11-169-232-62	Sequence 62, Appl
24	243.5	8.5	430	9	US-10-467-657-5700	Sequence 5700, Ap
25	243.5	8.5	430	9	US-10-467-657-7448	Sequence 7448, Ap
26	243.5	8.5	503	11	US-11-079-463-7133	Sequence 7133, Ap
27	243.5	8.5	1236	9	US-10-873-528-109	Sequence 109, App
28	241.5	8.5	758	11	US-11-096-568A-14593	Sequence 14593, A
29	235.5	8.3	375	11	US-11-096-568A-23618	Sequence 23618, A
30	235	8.2	1256	9	US-10-517-696-111	Sequence 111, App
31	233.5	8.2	766	11	US-11-096-568A-19864	Sequence 19864, A
32	232.5	8.2	1480	11	US-11-096-568A-28958	Sequence 28958, A
33	232	8.1	442	9	US-10-793-626-1000	Sequence 1000, Ap
34	231	8.1	1255	11	US-11-050-857-487	Sequence 487, App
35	231	8.1	1255	11	US-11-043-806-398	Sequence 398, App
36	230.5	8.1	429	11	US-11-079-463-7814	Sequence 7814, Ap
37	229.5	8.0	413	11	US-11-096-568A-20771	Sequence 20771, A
38	227.5	8.0	485	11	US-11-074-176-10	Sequence 10, Appl
39	227.5	8.0	715	11	US-11-072-512-3385	Sequence 3385, Ap
40	226.5	7.9	396	9	US-10-510-386-238	Sequence 238, App
41	226.5	7.9	399	9	US-10-510-386-30	Sequence 30, Appl
42	224.5	7.9	312	11	US-11-096-568A-12515	Sequence 12515, A
43	224	7.9	479	11	US-11-098-686-10838	Sequence 10838, A
44	222	7.8	535	11	US-11-096-568A-20039	Sequence 20039, A
45	221.5	7.8	453	11	US-11-096-568A-17494	Sequence 17494, A

ALIGNMENTS

RESULT 1
US-10-467-657-5462
; Sequence 5462, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5462
; LENGTH: 525
; TYPE: PR1
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5462

Alignment Scores:
Pred. No.: 1.79e-170
Score: 2681.00
Length: 525
Matches: 518
Percent Similarity: 99.8%
Conservative: 1
Best Local Similarity: 99.6%
Mismatches: 1

```
Query Match: 94.0% Indels: 0
DB: 9 Gaps: 0
US-10-665-990A-13 (1-1561) x US-10-467-657-5462 (1-525)

QY 2 AAAATACAGGCAATCCGCTCTGAATACTATATCCCGCATGAAACACGACGCTCATTTCC 61
DB 6 LysileGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer 25
QY 62 CTTTATATGCTCTCTCTGTTTCATGTTCTTTCATGTTGCCCGCCACTCGAAGAACGGACG 121
DB 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProProLeuGluGluArgThr 45
QY 122 GAAACCGCTCATTTCAATACATTCCAAACCTGTCCTCTGGACAAACATCCTGCAAAATCCGG 181
DB 46 GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTCATAACACGGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAGCC 241
DB 66 HisThrProHisAsnGlyLeuSerAspIleIleLeuLeuAspAsnIleLeuGlnIleArg 85
QY 242 CTTGCGCGCGCGCGCTTATCGAATCTGCGGAACACAGCCTCGATTGCAATACATAC 301
DB 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCGAACGACATTTCCGGCAGGCTGTGTTCAACCTCATGTACTTTCGCGCAGAA 361
DB 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaAlaGlu 125
QY 362 CGCGCGGTGGCGGTACGCTGTGTGGACGACAAACACGCGGGGTGGACGATCTC 421
DB 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTTCAACCCCTTCTGCTCA 481
DB 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATGGCGGCACTCGGCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGAC 541
DB 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMethis 185
QY 542 AACAAATCTTTTACCGCGACAAACCGCCACCATCTCGCGGACGCAATATCGCGAC 601
DB 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyArgAsnIleGlyAsp 205
QY 602 GAATATCTTCAAAGTCGTGAGACACCGTCTTTCGCGCACCTGGACATCTCGCCACCGC 661
DB 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTGTGGGGAAGTATCGCACGACTTCGACCGCTACTGGGCAAGCATTCGCCCCAC 721
DB 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCACGCGCATCATCGCAGCGGCAACATCGCAAGGTCTTCAAGCACTCGGATAC 781
DB 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGACGAAACATCCAGACACGCGCTCTCGCTACCGCGAACCGTCCGAACAGTCGCC 841
DB 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTTCTACCAAAAATACAGCGGAGCGCATCGACTGGCAGAGGTCCTCAACCGCGCTGATC 901
DB 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGACTCGACCGCGACCGCGCAACCGCCGATTCGCGGAGG 961
DB 306 SerAspSerProAlaLysGlyLeuAspArgAspArgLysProProIleAlaGlyArg 325
QY 962 CTGCAAGACGCGCTCAACACGCGCGGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC 1021
DB 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
```

```
QY 1022 CCTACAAATCCGCGCACAGACGCACTGGCAAACTGGTGCAGGCGCATAGACGTTACC 1081
DB 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCTGTGACCAACTCGCTACAGCGCACGACGTTGCGCGCTCCATTCGCGGTACGTCAAA 1141
DB 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAACCGCTGCTCAAAAGCCGCGCATCAAACTTACGAGCTGCAACCAACCATGCC 1201
DB 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
QY 1202 GTCCCGCGCAAAAGACAAAGCGCTGACCGCGAGCTCCGTAAACAGCTCGATGCCAAA 1261
DB 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTGGACGGCAAAACGCACTTTCATCGGCTCATTCACACCTCGACCCCGTTCC 1321
DB 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCACGGCTCAATACGAAATGGCGTGTCTATCGAAAGCCCCAAAATCGCAGAACAGATG 1381
DB 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGATACCAACCCGCAATACCGCTACCGGTTACCGTTCGACAAACAC 1441
DB 466 GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis 485
QY 1442 AACCGCTGCAATGGCACGATCCCGCCACCGCAAAACCTACCGCAACGAAACCGGAGCC 1501
DB 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAATTTGAAACGATCGCGCAAAATCTATCCCTGCTGCCCATCGAAGGTTTATTA 1561
DB 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu 525

RESULT 2
US-11-108-172-1068
; Sequence 1068, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
```

```
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1068

Alignment Scores:
Pred. No.:      8,48e-15      Length:      5179
Score:          336.50      Matches:      141
Percent Similarity: 33.6%      Conservative: 55
Best Local Similarity: 24.1%      Mismatches: 198
Query Match:      11.8%      Indels:      190
DB:               11      Gaps:        22

US-10-665-990A-13 (1-1561) x US-11-108-172-1068 (1-5179)
QY 119 ACGGAAGCGTCAATTTCAAT-----ACTTCCAAACCTGCTCTCGGAGAAC 166
Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTCTGCAAAATCCGGCACACCCCTCATACAAACGGGTATCCGACATCTACCTGCTCGAC 226
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
QY 227 GACCCCAAGAGCCCTTGGCGCGCGCGCCCTTATCGAATCTGCCAACACAGCGCTC 286
Db 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTGTGAATACTACATTTGGCGCAACGACATTTCCGGAGGCTGCTGTTCAACCTCA-- 344
Db 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACCTTGGCGAGAACGGCGGTGCGGTACGCC----- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
QY 381 -TGCTGTGG-----ACGACAACAAC 400
Db 1387 nCysCysTipProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
QY 401 ACGCGGGGTGACGATCTCTGCTCGCTCGACAGCCATCCCAATATCGAAGTGGCG 460
Db 1407 rProProProThrThrThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTTCGTCCTACGCAAAATGGCGCGACTCGGCTACCTGACCGACTCCCC 520
Db 1423 -----SerProProProThrThrThrThrThrProProProThrProSerProPr 1441
QY 521 CGCCTCAACCGCGCATGCACAAATCCTTTACCGCGCGACAAACCGCGCCACCATCTC 580
```

```
Db 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGGACGCAATATCGCGAGCAATACTTCAAGTCGGTAGGACACACCGTTTTCGCCGAC 640
Db 1457 -----ProIleSerThrThrThrProProThrThrThrThrProSerProProTh 1474
QY 641 CTGGACATCTCGCCACCGCAGCGTCTCGGCGAAGTATCGCACGACTTCGAC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
QY 696 -----GCTACTGGCAAGCCATTCCGCCCAACCGCCACGCGCATCATCCGCGGCGC 748
Db 1494 oProProThrThrThrProSerProMetThrThrProIleThrProProAlaSerTh 1514
QY 749 AAGATCGGCAAGGCTTTCAGGACTCGGATACAGCAACGAAACATCCAGACACGCGCTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGCTACCGCGAAACCGTCAAC-----AGTCGCCCTCTACCAAAAATACAGACG 862
Db 1529 rThrThrThrProProProThrThrThrProSerProProThrThr----- 1544
QY 863 GGACGCAATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCGCACACCCCTGCCAAAAGGA 922
Db 1545 -----ThrProIleThrProPr 1550
QY 923 CTGACGCGGACCGCGCAAAACCGCGATTTGCGGGAGGCTGCAAGACGCGCTCAACAG 982
Db 1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 983 CCGGAAAAAGGCTCATCTGTTTACCCCTATTTCGTCCTACAAATCCGGCACAGAC 1042
Db 1563 rProProProThrThrThrThrProProProThrThrThrProSerProProThrTh 1583
QY 1043 GCACCTGCAAAACTGGTGCAGGACGGCATAGAGCTTACCGTCTCTGACCA-----AC 1093
Db 1583 r-----ThrThrProSerProProThrIleThrTh 1593
QY 1094 TCCTACAGCGGACGACAGTTCGGCGTCCATTCGGGTACGTCAAAAT----- 1142
Db 1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 -----ACCGAAAAACCGTCTCAAGCCGCGCATCAAACTC 1177
Db 1613 oProThrThrThrProSerProProThrThrProIleThrProProThrSerThrTh 1633
QY 1178 TAGAGCTGCAACCCCAACCATGCGTCCCGCGCACAAAGACAAAGGCGCTGACCGGCGC 1237
Db 1633 rThrLeuProProThrThrThrProSerProProProThr-----ThrThrTh 1650
QY 1238 TCGGTAAACGACCTGATGCGCAAAACCTTCATTGTGGAGCGGCAACGCATCTTCATCGCG 1297
Db 1650 rProProProThrThrThrProSerProProThr-----ThrThrProSerProPr 1669
QY 1298 TCATTCAACTCGACCCCGTTCGCGACGCGCTCAATACCGAAATGGGCGTGCATCGAA 1357
Db 1669 eThrThrThrThrProProProThrThrThrProSerProIleThrThrThrPr 1689
QY 1358 AGCCCCAAAATCCGACAGACAGATGGAGCGCACCTCGCCGATACCAACCCGGAATACGCC 1417
Db 1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerPr 1708
QY 1418 TACCGGTTACCTCGACAAACACAAACCGCTGCAATGGCAGCATCCCGCCACCCGAAAA 1477
Db 1708 oIle-----ThrThrThrThrProSerSerThrThrThrProSerProPr 1725
QY 1478 ACCTACCGGAACGAAACCGGACCAAACTTTGGAACGCGATCGCCGCAAAAATCCTATCC 1537
Db 1725 oThrThrMetThrThrProSerProThrThrThrProSerProProThrThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
```

Db	1745	rThrLeuPro 1748	DB:	11	Gaps:	17
RESULT 3						
US-11-045-004-2518						
; Sequence 2518, Application US/11045004						
; Publication No. US20060078901A1						
; GENERAL INFORMATION:						
; APPLICANT: BUCHRIESER, CARMEN						
; APPLICANT: FRANGEUL, LIONEL						
; APPLICANT: COUVE, ELISABETH						
; APPLICANT: RUSNIOK, CHRISTOPHE						
; APPLICANT: FSIHI, HAFIDA						
; APPLICANT: DEHOUX, PIERRE						
; APPLICANT: DUSURGOT, OLIVIER						
; APPLICANT: CHETOUANI, FARID						
; APPLICANT: NEDJARI, HAFED						
; APPLICANT: GLASER, PHILIPPE						
; APPLICANT: KUNST, FRANCK						
; APPLICANT: COSSART, PASCALE						
; APPLICANT: DANIELS, JUSTIN						
; APPLICANT: GOEBEL, WERNER						
; APPLICANT: KREFT, JURGEN						
; APPLICANT: KUHN, MICHAEL						
; APPLICANT: NG, EVA						
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO						
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO						
; APPLICANT: GARRIDO-GARCIA, PATRICIA						
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO						
; APPLICANT: AMEND, ALEXANDRA						
; APPLICANT: CHAKRABORTY, TRINAD						
; APPLICANT: DOMANN, EUGEN						
; APPLICANT: HAIN, THORSTEN						
; APPLICANT: BERCHE, PATRICK						
; APPLICANT: CHARBIT, ALAIN						
; APPLICANT: DURANT, LIONEL						
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO						
; APPLICANT: BAQUERO, FERNANDO						
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO						
; APPLICANT: GOMEZ-LOPEZ, NURIA						
; APPLICANT: MADUENIO, ENCARN						
; APPLICANT: PABLOS, BETRIZ DE						
; APPLICANT: WEHLAND, JURGEN						
; APPLICANT: KARST, UWE						
; APPLICANT: ENTIAN, KARL-DIETER						
; APPLICANT: HAUF, JORG						
; APPLICANT: ROSE, MATTHIAS						
; APPLICANT: VOSS, HAMUT						
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES						
; FILE REFERENCE: 05394.0018-02						
; CURRENT APPLICATION NUMBER: US/11/045,004						
; CURRENT FILING DATE: 2005-01-28						
; PRIOR APPLICATION NUMBER: 10/637,657						
; PRIOR FILING DATE: 2003-08-11						
; PRIOR APPLICATION NUMBER: 10/257,023						
; PRIOR FILING DATE: 2002-10-08						
; PRIOR APPLICATION NUMBER: PCT/FR01/01118						
; PRIOR FILING DATE: 2001-04-11						
; PRIOR APPLICATION NUMBER: FR 00/04,629						
; PRIOR FILING DATE: 2000-04-11						
; NUMBER OF SEQ ID NOS: 2854						
; SOFTWARE: PatentIn version 3.3						
; SEQ ID NO 2518						
; LENGTH: 504						
; TYPE: PRT						
; ORGANISM: Listeria monocytogenes						
US-11-045-004-2518						
Alignment Scores:						
Pred. No.:			3.8e-12		Length:	
Score:			294.50		Matches:	
Percent Similarity:			37.7%		Conservative:	
Best Local Similarity:			23.5%		Mismatches:	
Query Match:			10.3%		Indels:	


```
QY 746 GGCACATCGCAGAGGTCTTCAGCA-----CTCGGATACAAGCAGAAACATCC 796
|||
Db 263 GlyLysAlaValTyrAlaMetGlnThrArgPheIleMetAspTrpAsnSerAlaSerSer 282
|||
QY 797 AGACACGGCTC-----CTGGCTACCGCGAAACCGTC-----GAACATCGGCC 841
|||
Db 283 ThrHisLysIleAspTyrLysAlaArgTyrPheProThrPheHisGlyLysGlyHisThr 302
|||
QY 842 CTCTACCAAAATACAGACGGGA---CGCATCGACTGGCAGAGCGTCCAAACCGCCTG 898
|||
Db 303 SerMetGlnIleValSerSerGlyProAspSerGluTrpGlnGlnIleLysAsnGlyTyr 322
|||
QY 899 ATCAGCGACCCCTCGCAAAAGGACTCGACCGGACCGCCGCAAAACCCGATGCCGGG 958
|||
Db 323 IleLys----- 324
QY 959 AGGTGCTCAAGACGGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTC 1018
|||
Db 325 -----MetIleAsnAlaLysLysThrIleTyrLeuGlnSerProTyrPhe 340
|||
QY 1019 GTCCCTCAAAATCCGGCAGACGACTGGCAAAACCTGGTGCAAGACGGCATAGAGTT 1078
|||
Db 341 IleProAspAlaSerLeuLeuGluAlaIleLysIleAlaAlaLeuSerGlyValAspVal 360
|||
QY 1079 ACCGTCTGACCAACTCGCTACAGCGGACCGACGCTTGGCGCGTCCATTCGGGTACGTC 1138
|||
Db 361 ArgValMetIleProAsn-----LysProAspHisAlaPheValTyrArgAlaThrThr 378
|||
QY 1139 AAATACCGGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCAT 1198
|||
Db 379 AsnTyrAlaGlyGluLeuMetGluThrGlyAlaLysIlePhe----- 392
|||
QY 1199 GCCGTCCCGCCACAAAAGACAAAGCGCTGACCGGACGCTCCGTAACAGCGCTGCATGCC 1258
|||
Db 393 -----IleTyrAspAsnGlyPhe-----IleHisAla 401
|||
QY 1259 AAACCTTCATTGTGGAGGCAAGCGCATCTTCATCGGCTCATTCACCTCGACCCCGT 1318
|||
Db 402 LysThrLeuValValAspGlyGluIleAlaSerValGlyThrAlaAsnMetAspPheArg 421
|||
QY 1319 TCCGACACGGCTCAATACCGAAATGGCGTGTGTCATCGAAAGCCCAAAATCGCAGAACAG 1378
|||
Db 422 SerPheArgLeuAsnPheGluValAsnAlaPheIleTyrGluLysGlnMetValGlnLys 441
|||
QY 1379 ATGGAGCGCACCTCGCCGAT-----ACCAACCCGAA---TAC 1414
|||
Db 442 LeuGluAspAlaPheLeuGluAspIleLeuLysSerTyrGlnLeuThrProGluLeuTyr 461
|||
QY 1415 GCCTACCGC 1423
|||
Db 462 AlaLysArg 464
|||
```

```
RESULT 5
; Sequence 1641, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1641
; LENGTH: 1970
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-10-821-234-1641
Alignment Scores:
Pred. No.: 3,79e-09 Length: 1970
Score: 250.50 Matches: 123
Percent Similarity: 38.2% Conservative: 69
Best Local Similarity: 24.5% Mismatches: 196
Query Match: 8.8% Indels: 115
DB: 9 Gaps: 24
US-10-665-990A-13 (1-1561) x US-10-821-234-1641 (1-1970)
QY 108 TGGAAAGACGGAGGAAGCGGTCAATTCAATATTTGCAAACTGTCTCTCTGGACAACA 167
|||
Db 1520 TrpAsnGlnGly-----AlaThrProAlaTyrGlyAlaTrpSerPro 1533
|||
QY 168 TCCTTGCAAAATCCGGCAGCACACCCCTCATAAACAGCGGTATCCGACATCTACCTGC- 221
|||
Db 1534 SerValGlySerGlyMetThrProGlyAlaAlaGlyPheSerProSerAlaSerAsp 1553
|||
QY 222 -----TCGACGACCCCCACGAGCCCTTG 245
|||
Db 1554 AlaSerGlyPheSerProGlyTyrSerProAlaTrpSerProThrProGlySerProGly 1573
|||
QY 246 CCGCGCGCGCGCTTATCGAATCTGCGGAACACAGCCTCGATTGCAATACATATT 305
|||
Db 1574 SerProGlyProSerSerProTyrIleProSer----- 1584
|||
QY 306 GGCACACGACATTTCCGGCAGCGTGTGTTCAACCTCATCTACCTTCCCGCAGAACCG 365
|||
Db 1585 -----ProGlyGlyAlaMetSerProSerTyrSer--ProThrSerPro 1598
|||
QY 366 GCGTGGCGGTACGCTGCTGTGGACGACAAACACGCGCGGGTGTGACCATCTCCTGC 425
|||
Db 1599 Ala-----TyrGluProArgSerProGlyGlyTyrThrProGln-- 1611
|||
QY 426 TCGCCCTCGACAGCCATCCCAATATCGAAGTGGCCCTGTTCAACCCCTTCGTCTTACGCA 485
|||
Db 1612 SerProSerTyrSer--ProThrSerProSerTyrSerProThrSerProSerTyrSer 1630
|||
QY 486 -----AATGGCGGCACTCGGCTACCTGACCGACT--TCCCGCGCTCAACCGCCGCA 536
|||
Db 1631 ProThrSerProAsnTyrSerProThrSerProSerTyrSerProThrSerProSerTyr 1650
|||
QY 537 TGCACAAATAATCTTTACCGCCGACACACCGCCGCCCATCTACTCGCGGAGCGCAATATCG 596
|||
Db 1651 SerProThrSerProSerTyrSerProThrSerProSerTyrSerProThrSerProSer 1670
|||
QY 597 GCGACGAATCTTCAAAGTCGGTGAGGACACCGTTTTTCGCGGACCTGGACATCTCTCGCA 656
|||
Db 1671 TyrSerProThrSerProSerTyr--SerProThrSerProSerTyrSerProThrSer 1689
|||
QY 657 CCGCAGCGCTCG-----TCGGCGAAGTATCGCACGACTTCGACCGTACTGGGCAAGCC 710
|||
Db 1690 ProSerTyrSerProThrSerProSerTyrSerProThrSerProSerTyrSerProThr 1709
|||
QY 711 ATTCCGCCCAACAGCCACGCGCATCATCCGACGGGCAACATCGGCAAGGGTCTTCAAG 770
|||
Db 1710 SerProSerTyrSerProThrSerProSerTyrSerProThrSerPro--Ser 1726
|||
QY 771 CACTCGGATACACGACGAAACATCCAGACACGCGCTCTCGGCTACCGCGAAACCGTCG 830
|||
Db 1727 TyrSerProThrSerProSerTyrSerProThrSerProSerTyrSerProThrSerPro 1745
|||
QY 831 AACAGTCCGCCCTCTACCAAAAAATACAGCGGACGCGCATCGACTGGCAGAGCGTCCAAA 890
|||
Db 1747 AsnTyrSerProThrSerProAsnTyrThrProThrSerProSerTyrSerProThrSer 1766
|||
QY 891 CCGCGCTGATCAGGACACACCCCTCGAAAGACTCGACCGGACCGCGCAACCCGCA 950
|||
Db 1767 Pro-----SerTyrSerProThrSerProAsnTyrThrProThrSerProAsn----- 1782
|||
QY 951 TTGCGGGAGGCTGCAAGACGCGCTCAAAACAGCCGCGCTCTATCTGTTTTCAC 1010
|||
```

Db 1783 -----TyrSerProThrSerProSer-----TyrSer 1791
Qy 1011 CCTATTTCGTCTACAAATCCGGCAGACGACTGGCAAACTGTGTGAGGACGCA 1070
Db 1792 ProThrSerProSerTyrSerProThrSerProSerTyrSer----- 1805
Qy 1071 TAGAGCTTACCGTCTCGACCACTCGCTAC-----AGCGACCGAGCTGTGCGCGCTCC 1124
Db 1806 -----ProSerSerPro---ArgTyrThrProGlnSerProThrTyrThrProSer 1821
Qy 1125 ATTCGGCTACGTCAAAATACCGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGC 1184
Db 1822 SerProSerTyrSer-----ProSerSerProSerTyrSer 1833
Qy 1185 TGCAACCCACCATCGCTCCCGCCGACAAAGACAAAGGCTGACCGGAGCTCGTAA 1244
Db 1834 -----ProThrSerProLysTyrThr-----ProThrSerPro--- 1844
Qy 1245 CCAGCTGCATGCCAAACCTTCATTGTGGACGGCAAAACGCATCTTCATCGGCTCATCA 1304
Db 1845 -----SerTyrSer 1847
Qy 1305 ACTTCGACCCGTTCCGCGAGGCTCAATACGAAATGGCG---TCGTCAATCGAAAGCC 1361
Db 1848 ProSerSerProGluTyrThrProThrSerProLysTyrSerProThrSerProLysTyr 1867
Qy 1362 CCAAAATCGCAGACAGATGAGCGCACCTCGCGATACACACCGAATACGCTACC 1421
Db 1868 SerProThrSerProLysTyrSerProThrSerProThrTyrSerProThrProLys 1887
Qy 1422 GCGTTACCTCGCAAAACACACCGCCCTGCAATGGCAGATCCCGCCACCGGAAACCT 1481
Db 1888 TyrSerProThrSerProThrTyrSerProThrSerProValTyrThrProThrSerPro 1907
Qy 1482 ACCCGAAGCAGCCGAGACCAAACTTTGGAAACGATCCCGCAAAATCTATCCCTGC 1541
Db 1908 LysTyrSerProThrSerProThrTyrSerProThrSerProLysTyrSer---ProThr 1926
Qy 1542 TGCCCATCG 1550
Db 1927 SerProThr 1929

RESULT 6
US-11-096-568A-14646
; Sequence 14646, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14646
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(702)
; OTHER INFORMATION: Ceres Seq. ID no. 11049476
US-11-096-568A-14646

Alignment Scores:
Pred. No.: 4,15e-09 Length: 702
Score: 249.00 Matches: 159
Percent Similarity: 30.6% Conservative: 38
Best Local Similarity: 24.7% Mismatches: 213
Query Match: 8.7% Indels: 234
DB: 11 Gaps: 30

US-10-665-990A-13 (1-1561) x US-11-096-568A-14646 (1-702)
Qy 73 CTTCTCTGTTTCATGTTCTTTCATGTTGCCCCCACTGGAAGAACGGACGGAAGCCGTCA 132
Db 38 ProAlaHisLeuCysPheHisThrProAlaProAlaGlyProValGlnGly---ProGly 56
Qy 133 TTTCAATACTTCCAAACG-----TGCTCT 156
Db 57 LeuGlnPheArgGlnValLeuAlaAlaArgArgGlnLeuArgAlaAlaArgValPro 76
Qy 157 CTTGGACAACTCTGCAATCCGGCACACCCC-----TCATAACAACGGGCTATCCGA 210
Db 77 ProGlyAlaValProGlnAlaArgArgProAlaCysAlaValGlnGluAlaArgArg 96
Qy 211 CATCTACTGCTCGACGACCCCGACGAGCCCTTGGCGCGCGCGCGCCCTTATCGAATC 270
Db 97 His-----CysArgProArg----- 101
Qy 271 TGCCGAACACACGCTCGATTGTGCAATACTACATTTGGCGCAACGACATTTCCGGCAGGCT 330
Db 102 -----AlaArgGlyAlaGluGluSerGlyGlyArgGlnGlu 114
Qy 331 GCTGTTCAACCTCATGTACCTTGGCAGAACGGCGTGCCTG----- 375
Db 115 AlaAlaGln---LeuValProAlaArgArgGlyGlyArgAlaArgValArgGlyArg 133
Qy 376 -----ACGCTGCTGTTGGACGACAAACACGCGCGGGTGGACGA 417
Db 134 ValArgGlyGlnArgProArgAlaAlaProArgAlaAlaGlnHisGlyArgProGlyArg 153
Qy 418 TCTCCTGCTGCTCGACGACCATCCCAATATCGAAGTGGCTGCTTCAACCCCTCGT 477
Db 154 AlaAspArgArg-----GlnLeuArgAla-----Arg 162
Qy 478 CTTAGCGAAATGGCGCAGCTCTGCTACCTGACCGACTTCCCGCCTCAACCCGCGAT 537
Db 163 ArgGlyGluAlaAlaArgAlaPro---ArgLeuArgProHisProAlaLeuHis 181
Qy 538 GCACAAACAATC---CTTTACCGCGGACAAACG----- 567
Db 182 ArgArgGlyLeuArgAlaGlnArgArgGlnProLeuLeuArgProArgArgGlyAlaPro 201
Qy 568 -----CGCCACCATCTCTCGGCGG-----ACGCAATATCGGCGCAATA 606
Db 202 ProLeuArgArgArgArgHisGlyAlaArgAlaAlaAspThrGlnArgArgArgVal 221
Qy 607 CTTCAAAATCGGTGAGGACACCGGTTTTCGCGGACCTGACATCTCGCACCGGAGCGT 666
Db 222 LeuArgLeu-----ProLeuHisGlyAspArg 230
Qy 667 CGTCGGGAAGTATCGCAGCAGCTTCGACCGCTACTGGGCAAGCCATTCGCGCCACACGC 726
Db 231 AlaProAlaProGlyAlaArgAspArgPro-----ProArgLeuProGlnGly 246
Qy 727 CACGCGCATCATCCGCGAGCGCAACATCGCGAAGGGTCTTCAAGCACTCGGATACACGA 786
Db 247 HisTrpArg-----AlaAla 251
Qy 787 CGAAACATCCAGACACCGCTCTCTGCGCTACCGGGAACCGTCGAAAGTCCGCCCTCTA 846
Db 252 ArgAlaLeuArgAspArgAlaProArgProValLeuArgAlaArgArgValArg----- 269
Qy 847 CCAAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCGCCTGATCAGCGA 906
Db 270 ProArgGlyProLeuGlyArgHisArgGlyGlnArgArgProArgLeuLeuGlnArg 289
Qy 907 CACCCCTGCAAAAGGACTCGACCGGACCGCAAAACCGCCGATTTCGCGAGGCTGCA 966
Db 290 ProProAlaArgHisValArgGlyArgProProLeuProArgPro---ArgGlyGlyArg 308
Qy 967 AGACGCGCTCAACAGCCCGGAAAAAAGCGTCTATCTGGTTTCAACCTATTTCTGCTCCCTAC 1026
Db 309 ArgArgProArgAspArgArgProLeuArgLeuArg-----ArgAlaAla 323

```
QY 1027 AAAATCCGACAGACGACTGCGCAAAACTGCTGACGAG- 1065
Db 324 ArgLeuArgHisAspArgAlaProGlnAlaGlyProGlyHisArgGlyAlaProArgAla 343
QY 1066 -----CGCATAGAGTTACGCTCTGACCAACTCGCTACAGGCGACGAGTTGC 1116
Db 344 GlnLeuArgArgHisGlnGlnAlaValProGlnValLeuLeuGlnAlaArg----- 361
QY 1117 CGCCGCTCCATTCGGCTACGTCAAAATACCGAAACCGCTGCTCAAAAGC----- 1164
Db 362 ArgHisGlnValArgArgArgGlyAspProAlaGlyAlaAlaHisAspAspProArgLeu 381
QY 1165 -----CGGCATCAAACTTACGAGCTGCAACCCCAACCATGCGCTCC----- 1206
Db 382 ArgHisArgGlnLeuArgGlyCysAlaArgPro-----ProGlyGlyValGlnAla 399
QY 1207 -----CGCCAAAGACAAAGCGCTGACCGGAGCTC 1239
Db 400 ProGlyAspAlaAlaArgArgValAlaArgGlyAlaGlyGlnGlyGluAspValAlaVal 419
QY 1240 CGTAAACAGCTGCTGCAAAAC----- 1263
Db 420 ArgArgAlaProGlnAlaArgArgGlyArgValGlyAspGlyValGlyArgAlaGly 439
QY 1263 ----- 1263
Db 440 LeuLeuLeuProProValGluArgValGlyArgGlyAspGlyArgGlyGlyGly 459
QY 1264 -----CTTCATTGTGGCGCAACGCACTTCATCGGCTCATTCAA----- 1305
Db 460 AspArgLeuLeuHisAspProArgArgLeuHisLeuGlnArgValArgAlaProGly 479
QY 1306 -----CCTCGACCC 1314
Db 480 GluArgAlaAspArgAspProGlyArgAlaHisGlyProValHisAlaProArgArg 499
QY 1315 CGGTTCCGACCGCTCAATACCGAAATGGCGCTGCTATCGAAAGCCCAAAAT----- 1368
Db 500 ProAlaAlaValAlaGlyGluProGlyGlyGlyHisGlyGluProGlnProAlaGly 519
QY 1368 ----- 1368
Db 520 ProArgGluProValArgValProArgGlyGlyGlyAlaValAlaGlnGlyValGlyLeu 539
QY 1369 CGCAGAACAGATGGAGCGCACCTCGCCGATACCAACCGCAATACGCTACCGGTTAC 1428
Db 540 ArgGlnGlyGlyProValHisGlyArgAlaHisGlnValArgValArg-----ArgGly 557
QY 1429 CCTCGACAAACACACCGCTGCAATGGCAGCATCCGCGCACCCGAAACCTACCCGAA 1488
Db 558 ProValArgArgArg-----AlaLeuLeuArgSerHisGlyProGlyArgGlyProPro 575
QY 1489 CGAACCCGAGGCAAACTTGGAAACGATCCCGCAAAATCCTATCCCTGCTGCCAT 1548
Db 576 AlaArgArgGlyArgValArgAlaHisLeuArg-----ProArgArgAla 591
QY 1549 CGAAGCTTTATT 1560
Db 592 ArgArgHisVal 595
```

```
RESULT 7
US-11-096-568A-11657
; Sequence 11657, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
```

```
; SEQ ID NO 11657
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(493)
; OTHER INFORMATION: Ceres Seq. ID no. 15219800
US-11-096-568A-11657

Alignment Scores:
Pred. No.: 4,61e-09 Length: 493
Score: 248.00 Matches: 126
Percent Similarity: 35.5% Conservative: 61
Best Local Similarity: 23.9% Mismatches: 163
Query Match: 8.7% Indels: 178
DB: 11 Gaps: 18

US-10-665-990A-13 (1-1561) x US-11-096-568A-11657 (1-493)
QY 34 CCCGATGAAACACGACGCTCATTTCCCTTTTATGCCTCTCTCTGTTTCATGTTCTTC 93
Db 28 ProAspProSerAlaSerProProProProProProProProProProProProProPro 47
QY 94 ATGGTTGCCCTGAGGAAAGACGCGAAG---CCGTATTTCATCTACTTCCAAACC 150
Db 48 AlaSerSerProSerThrArgThrAlaGlyAlaLeuProThrCysArgAlaArgArg 67
QY 151 TGTCTCTCTGGACAACATCCTGCAATCCGCGCACACCCCTCATAAACGGGCTATCCGA 210
Db 68 CysTrpAlaGly----- 71
QY 211 CATCTACTGTCGACGACCCCAAGCCCTTGGCCCGCGCGCCCTTATCGAATC 270
Db 72 -----ProSerArgArg---ProArgSerProSerArgProSerProAlaProThrSer 88
QY 271 TGCGAACACAGCTCGATTTGCAATACTACATTTGGCGCAACGACATTTCCGGCAGGCT 330
Db 89 TrpAlaAlaSerProThrGlyArgAlaThrArg---AlaThrArgAlaArgThrAlaThr 107
QY 331 GCTGTTCAACCTCATGTACCTTGGCGCAGAACGCGCGTACGCTGCTCTGTTGA 390
Db 108 ThrAlaArgProProProAlaThrArgSerSerThrSerProArgProThrThr--- 126
QY 391 CGAACAAACACGCGCGGTTGGACGATCTCTGCTCGCCCTCGACAGCATCCCAATAT 450
Db 127 -----ArgSerProProProProProProProProProProProProProSerAla 138
QY 451 CGAAGTGCCTGTTCAACC---CCTTCGTCCTACGCAATGGCGCGCACTCGGCTACCTGA 509
Db 139 SerSerThrProSerArgProSerThrArgAlaSerGlyProSerGlyAlaSer--- 157
QY 510 CGGACTTCCCGCTCAACCGCATGCGCAACAAATCCTTTACCGCGCAACACCGCG 569
Db 158 ThrSerGlyProSerSerProSerAlaAlaThrArgArgSerArgProSerAlaAlaArg 177
QY 570 CCACCATACTCGGCGGACGCAATATCGCGGACGAAAT----- 605
Db 178 ProArgArgSerAlaProAlaSerAlaThrThrArgThrThrAlaProThrThrThr 197
QY 606 -----ACTTCAAAGTCGTTGAGGACACCGTTTTCGCGCACCTGCGACATCTCGCCACCG 659
Db 198 ArgAlaSerAlaAlaThrSerArgThrProProProProProProProProProProThr 217
QY 660 GCAGCGCTCGTGGCGAAGTATCGCACGACTTCGACCGCTACTGCGGCAAGCCATTCGCGCC 719
Db 218 GlyProSerSerSerArgSerProSerProThrSerProSerProSerProSerProSerPro 237
QY 720 ACAACGCGACGCGCATCATTCGCGACGCGCAACATCGGCAAGGCTCTTCAAGCACTCGGAT 779
Db 238 ThrSerProArgThrCysSerSerAlaAlaProSerThrArg----- 251
QY 780 ACAACGACGAAACATCCAGACACGCGCTCTCGCGTACCGCGAAACCGTCGAAACAGTCGC 839
```

```
Db 252 -----ThrThrAlaProThrThrAlaSer-----ThrProArgArg 263
QY 840 CCCTCTACCAAAATACAGCGGAGCGACTGCGAGCGTCCAAACCCGCTGA 899
Db 264 ProAlaAlaSerSerAlaSerArgThrAlaSerSerArgSerProPro 281
QY 900 TCAGCGACACCCCTGC AAAAGAGACTCGACCGCGCGCGCAAAACCGCGATTGCGCGGA 959
Db 282 ---ThrThrPro----- 284
QY 960 GGCTGCAAGACGGCTCAACAGCCGGAAGAGCGTCTATCTGTTTTCACCTATTTCG 1019
Db 285 -----Ser 285
QY 1020 TCCCTACAAATCCGGGCACAGACGCACTGGCAAAACTGGTGCAAGCGCATAGACGTTA 1079
Db 286 SerAlaAlaSerPro----- 290
QY 1080 CCGTCTGACCAACTCGCTACAGCGCGACGAGTTCGCGCGTTCATTCCGGCTACGTCA 1139
Db 291 -----ArgArgThrArgProProSerSerProProPro 302
QY 1140 AATACCGAAACCGCTGCTCAAGCGCGCATCAACTCTACGAGCTGCAACCCCAACCATG 1199
Db 303 SerSerProProSerCysAlaArgProAlaAlaSerSerProGly----- 317
QY 1200 CCGTCCCGCCACAAAGACAAAGCGCTGACCGGAGCTCGTAACAGCGCTGCATGCCA 1259
Db 317 ----- 317
QY 1260 AAACCTTCATTGTGAGCGGCAAGCGCACTTCATCGGCTCATTAACCTCGACCCCGGTT 1319
Db 318 -----ThrLeuSerCysSerAlaSerAlaThrSerSerSer----- 329
QY 1320 CCGCAGCGCTCAATACCGAAATGGCGTCTCATCGAAGCGCCCAAAATCG----- 1370
Db 330 -----SerThrSerAlaArgAlaProSerSerIleCysSer 341
QY 1371 ---CAGAACAGATGAGCGCACCTCGCGGATACACACCGGAATA---CGCTACCGCG 1424
Db 342 LeuSerThrArgLeuArgSerSerSerSerLeuArgThrArgMetThrSerIleProAr 361
QY 1425 TTACCCTCGACAAACACACACCGCTGCAATGGCAGCATCCCGCCACCGCAAAACCTACC 1484
Db 361 gArgProLeuLeuSerArgProProThrSerThrArgThr----- 374
QY 1485 GGAAGACCGCGAAGCGCAACTTTGG-----AAACGATCGCGCGCAAAATCC 1532
Db 375 -----SerArgSer-ArgCysTrpCysValMetAlaArgArgLeuProLeuMetSerL 392
QY 1533 TATCCCTGCTGCCCATC 1549
Db 392 euthrArgLeuProLeu 397

RESULT 8
US-11-169-232-14
; Sequence 14, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
```

```
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseri, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 941 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-11-169-232-14
Alignment Scores:
Pred. No.: 5,42e-09 Length: 941
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-14 (1-941)
QY 144 CCAAACTGCTCTCTCGACAAACATCTGCAAAATCCGGCACACCCCTCATAACACGGGC 203
Db 132 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 149
QY 204 TATCCGACATCTACC-----TGCTCGACGAGCCCGCCAGAGCCCTTGCGC 248
Db 150 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 169
QY 249 -----CCGCGCGCGCTTATCGAATCTGCGGAACAGACGCTCGATT 290
Db 170 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 189
QY 291 TGAATACTACA-----TTTGGCGCAACGACATTT 320
Db 190 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 209
QY 321 CCGGAGGCTGCTTCAACCTCATCTTCCCGCAGAACGCGGCGTGGCGGTACGCC 380
Db 210 ProLysGluProAlaProThrThrThrLysGluProAla----- 222
QY 381 TGCTGTTGGACGACAAACACACGCGCGGTTGACAGATCTCTGCTCGCCCTCGACAGCC 440
Db 223 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 239
QY 441 ATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCCTTACGCAATGGCGCGCATCG 500
```

Db 240 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 258
QY 501 GCTACTGACACTTCCCGCCCTCAACCGCGCATGCACAAACAAATCCTTTACCG--- 557
Db 259 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 278
QY 558 -----CCGCAAC----- 566
Db 279 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 298
QY 567 ---GGCCACCATCTCGCGGACGAATATCGCGGAGAAATCTTCAAAGTCGGTGAGG 623
Db 299 ProAlaProThrThrProLysGluProAlaProThrThrLysGluProSerProThr 318
QY 624 ACACGGTTT-----TCGCGGACCTGGACATCTCGCCACCGACGCGTCGTCGGCGAAG 677
Db 319 ThrProLysGluProAlaProThrThrLysSerAlaProThrThrLysGluPro 338
QY 678 TATCGACGACTTCGACCGCTACTGGGCAAGCATTCCGCCACAAACGCCGCGCATCA 737
Db 339 AlaProThrThrLysSerAla-----ProThrThrProLysGluPro 353
QY 738 TCCGAGCGGCAACATCGCAAGGCTCTTCAAGCACTCGGATACAAACGACGAAACATCA 797
Db 354 SerProThrThrLys-----GluProAlaProThrThrProLysGluPro 369
QY 798 -----GACACGGCTCTCGCTACCGGAAACCGTCAACAGTCGCCCTCTACC 848
Db 370 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 389
QY 849 AAAAAATACAGAGGCATCGACTGGCAGAGGTCCAAACCGCGCTGATCAGCG--- 905
Db 390 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 406
QY 906 -----ACACCCCTGCAAAAGGACTCGACGCGACCGCGCAACCGC 947
Db 407 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLysLys 424
QY 948 CGATTGCGGGAGGTGCAAGACGGCTCAAAACGCGGAAAGCGTCTATCTGGTTT 1007
Db 425 LeuThrPro-----ThrThrProLysLysLeuAla----- 434
QY 1008 CACCCTATTCCTCGCTACAAATCGGACAGAGCACTGCAAAACTGGTGCAGGAGC 1067
Db 435 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 449
QY 1068 GCATAGACGTTCGCTGACCAACTCGCTACAGCGACCGACGTCGCG----- 1118
Db 450 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 465
QY 1119 ---CCGTCAATTCGGCTAGCTCAATACCGAAACCGGTGCTCAAAGCGGCATCAAC 1175
Db 466 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 482
QY 1176 TCTACGAGTGCACCAACCATCGCTCGCCCGCACAAAGACAAAGCCCTGACGGCA 1235
Db 483 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 495
QY 1236 GCTCGCTAACCGCTGCTGCAAAACCTTCATTGTCGACGCAAAACGATCTTCATCG 1295
Db 496 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 514
QY 1296 GTCTATTCACTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGTGCATCG 1355
Db 515 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 528
QY 1356 AAGCCCAAAATCGCAGACAGATGCGAGCGCACCTCGCCGATACCACCGCAATACG 1415
Db 529 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 544
QY 1416 CTATCGCGCTTACCTCGCAACACACCGCGCTGCAATGGCAGCATCCCG---CCACCC 1472
Db 545 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 563

QY 1473 GAAAAACCTACCGAAGCAACCGGAAG 1499
Db 564 LysGluProAlaProThrThrProLys 572
RESULT 9
US-11-169-232-84
; Sequence 84, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserti, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1022 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-11-169-232-84
Alignment Scores: 5,48e-09 Length: 1022
Pred. No.: 247.50 Matches: 135
Score: 35.2% Conservative: 44
Percent Similarity: 26.5% Mismatches: 205
Best Local Similarity: 8.7% Indels: 125
Query Match: 11 Gaps: 29
DB: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-84 (1-1022)

QY 144 CCAAACCTGTCTCTGGACAAACATCTCGCAATCCGCAACCCCTCATACACACGGGC 203
Db 213 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 230

QY 204 TATCGACATCTACC-----TGCTGACGACGCCCCACGAAGCCCTTGCG 248
|||||
Db 231 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 250
-----CCCGCGCGCCCTTATCGAATCGCCGACACACGACCTCGATT 290
249 |||
Db 251 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 270
251 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
|||||
Db 271 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 290
321 CGGCGAGGCTGCTGTTCACCTCATGTACTCTGCGCGAGAACGCGCGCTGCGACGCC 380
|||||
Db 291 ProLysGluProAlaProThrThrThrLysGluProAla----- 303
381 TGCTTTGGACAGACAAACAGCGCGGTGGAGCATCTCTGCTCCCTTCGACAGCC 440
|||||
Db 304 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThrThr 320
441 ATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCTGCTACGCGAATGGCGGCACTCG 500
321 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 339
501 GCTACTCTGACGACTTCCCGCGCTCAACCGCGCATGCACAAACAAATCCTTTACCG--- 557
340 GluProThrProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThr 359
558 -----CCGACAAC----- 566
360 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 379
567 ---GCGCCACCATCTCGCGGAGCGCAATATCGGCGAGCAATACTTCAAAGTCGGTGAGG 623
380 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 399
624 ACACCGTTT-----TCGCCGACTTGGACATCTCCGCCACCGCAGCGCTCGCGCAAG 677
400 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 419
678 TATCGCAGACTTGCACCGCTACTGGGCAAGCCATTCGCCCAACAGCCGCGCATCA 737
420 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 434
738 TCCGAGCGCAACATCGCAAGGCTCTCAAGCACTCGGATACAGCAGCAACATCCA 797
435 SerProThrThrThrLys-----GluProAlaProThrThrThrProLysGluPro 450
798 -----GACACGCGCTCTGCGCTACCGCGAAACCGTCGAACAGTCGCGCCCTTACC 848
451 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 470
849 AAAAAATACAGCGGAGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCG--- 905
471 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 487
906 -----ACACCCCTGCAAAAGCACTCGACCGCGCGCGCGCAACCGC 947
488 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLysLys 505
948 CGATTGCGGAGGCTCGAAGCGCTCAACAGCGCCGAAAGCGTCTATCTGCTTT 1007
506 LeuThrPro-----ThrThrProGluLysLeuAla----- 515
1008 CACCTATTTCCTCCCTACAAAATCCGGCAGACAGCACTGGCAAAACTGGTGAGGAGC 1067
516 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 530
1068 GCATAGAGTTTACCGTCTCGCAACTCGTACAGGCGACGAGTTCGCG----- 1118
531 AlaProThrThrProGluGluProThr-----ProThrThrProGluGluPro 546

QY 1119 ---CGTCCATTCCGGTACGTCAAATACCGAAACCGCTGCTCAAACCGCGCATCAAC 1175
|||||
Db 547 AlaProThrThrProLysAlaAla---AlaProSerThrProLysGluProAla----- 563
1176 TCTACGAGCTCAACCAACCATCGCTCCCGCCGACAAAAGACAAAGGCGCTGACCGGCA 1235
564 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 576
1236 GCTCGTAACAGCGCTGCATCCGAAACCTTCATTGTGGAGCGGCAAGCAATCTTCATCG 1295
577 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 595
1296 GCTCATTAACACTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGCTGCTCATCG 1355
596 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 609
1356 AAAGCCCAAAATCGCAGACAGATGGAGCGCACCTCGCGATACACACCCGAATACG 1415
610 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 625
1416 CTTACCGGTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCG---CCACCC 1472
626 ThrThrSerAspLysProAlaProThrPro-----LysGlyThrAlaProThrThrPro 644
1473 GAAAAACCTACCGCAACGACCCGAAG 1499
645 LysGluProAlaProThrThrProLys 653
RESULT 10
US-11-169-232-74
; Sequence 74, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170

; TELEFAX: (617) 876-5851

; INFORMATION FOR SEQ ID NO: 74:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1038 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 74:

US-11-169-232-74

Alignment Scores:

Pred. No.: 5 5e-09 Length: 1038

Score: 247.50 Matches: 135

Percent Similarity: 35.2% Conservative: 44

Best Local Similarity: 26.5% Mismatches: 205

Query Match: 8.7% Indels: 125

DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-74 (1-1038)

QY 144 CCAAACTGTCCTCTGGACAACTCTGCAATCCGGCAGACCCCTCATCAACAGCGC 203

Db 197 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 214

QY 204 TATCCGACATCTACC-----TCCTGACGACCCCGACGAGCCCTTGCG 248

Db 215 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 234

QY 249 -----CCCCGCGCCCTTATCGAATCTCCGAAACACAGCCCTCGATT 290

Db 235 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 254

QY 291 TGCATATCTACA-----TTTGGCGCAACGACATTT 320

Db 255 AlaProThrThrProLysGluProAlaProThrThrLysGluProAlaProThrThr 274

QY 321 CCGGAGGCTGTCTCACTCATGTACCTTGGCGCAGAAACGGCGTGGCGTACGCC 380

Db 275 ProLysGluProAlaProThrThrThrLysGluProAla----- 287

QY 381 TGCTGTGGACGACCAACACGCGGGTGGAGCATCTCTGCTCGCCCTCGACAGCC 440

Db 288 -----ProThrThrLysSerAlaProThrThrProLysGluProAlaProThr 304

QY 441 ATCCCAATTCGAAGTGGCCCTGTTCACCCCTTCTGCTACGCAAAATGGCGCACTCG 500

Db 305 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 323

QY 501 GCTACTGACGACATCCCGCCCTCAACCGCGCATGCAACAAATCTTTACCG--- 557

Db 324 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 343

QY 558 -----CCGACAAACC----- 566

Db 344 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 363

QY 567 ---GGCCACCATATCGCGGAGCGCAATATCGGCGCAAGAACTTCAAAGTCGTGAGG 623

Db 364 ProAlaProThrThrProLysGluProAlaProThrThrLysGluProSerProThr 383

QY 624 ACACCGTTT-----TCGCGGACCTGGACATCTCGCCACCGGACGCGTGGCGGAAG 677

Db 384 ThrProLysGluProAlaProThrThrLysSerAlaProThrThrThrLysGluPro 403

QY 678 TATCGCAGCACTTCGACCCCTACTGGCAAGCCATTCCGCCCAACGCGCAGCATCA 737

Db 404 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 418

QY 738 TCCGAGCGGCAACATCGGCAAGGGTCTTCAAGCACTCGGATACAAACGAGAAACATCA 797

Db 419 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 434

QY 798 -----GACACGCGCTCTCGGCTACCGCGAAACCGTCGACAGTCGCGCCCTCTACC 848

Db 435 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 454

QY 849 AAAAAATACAGACGGGACGTCATCGATCGGTCAGAGCGTCAAAACCGCGCTGATCAGCG--- 905

Db 455 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 471

QY 906 -----ACACCCCTGCAAAAGACTCGACCGGACCGCGCAAAACCGC 947

Db 472 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLysLys 489

QY 948 CGATTGCGGAGGCTGCAAGACGCGCTCAAAACGCGGCAAAAGCGCTCTATCTGGTTT 1007

Db 490 LeuThrPro-----ThrThrProGluLysLeuAla----- 499

QY 1008 CACCTATTTCGTCCTACAAAATCCGGCAGACGCACTGGCAAACTGGTGCAGGACG 1067

Db 500 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 514

QY 1068 GCATAGAGTTACGTCCTGACCAACTCGTACAGGCGACGAGCTTCCG----- 1118

Db 515 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 530

QY 1119 ---CGTCCATTCCGGCTACGTCAAATACCGAAACCGCTCTCAAAACCGCGCATCAAAC 1175

Db 531 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 547

QY 1176 TCTACGAGCTGCAACCAACCATCGCTCCCGCCGCAACAAAGACAAAGCGCTGACCGCA 1235

Db 548 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 560

QY 1236 GCTCGTTAACAGCTGTCATCGCAAAACCTTCATTGTGGAGCGGCAACCGCATCTTCATCG 1295

Db 561 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 579

QY 1296 GCTCATTCACCTCGACCCCGTTCCGACCGCTCAATACCGAAATGGCGTGGTCACTCG 1355

Db 580 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 593

QY 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGCATACCAACCCGAATACG 1415

Db 594 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 609

QY 1416 CTACCGCGTTACCTCGCAACACACACCGCTGCAATGGCAGCATCCG---CCACCC 1472

Db 610 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 628

QY 1473 GAAAAACCTACCCGCAAGAACCCGAAG 1499

Db 629 LysGluProAlaProThrThrProLys 637

RESULT 11

US-11-169-232-58

; Sequence 58, Application US/11169232

; Publication No. US20060025570A1

; GENERAL INFORMATION:

; APPLICANT: Turner, Katherine

; Clark, Stephen C.

; Jacobs, Kenneth

; Hewick, Rodney M.

; Geisner, Thomas G.

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors

; NUMBER OF SEQUENCES: 143

; CORRESPONDENCE ADDRESS:

; ADDRESS: Genetics Institute, Inc.

; STREET: 87 Bridgepark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-Jan-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-Jun-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-Dec-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-Aug-1989

ATTORNEY/AGENT INFORMATION:

NAME: Cseri, Luann

REGISTRATION NUMBER: 31,822

REFERENCE/DOCKET NUMBER: GI 5190

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)876-1170

TELEFAX: (617)876-5851

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 1049 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 58:

US-11-169-232-58

Alignment Scores:

Pred. No.: 5.5e-09 Length: 1049
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-58 (1-1049)

```
Qy 144 CCAAACTGCTCCTCGTGAACATCTGCAAAATCCGGCACACCCCTCAATAACACGGCG 203
Db 240 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 257
Qy 204 TATCCGACATCTAC-----TGCTCGACGACCCCAAGAGCCCTTCGCG 248
Db 258 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 277
Qy 249 -----CCGCGCGCGCTTATCGAATCTGCGGACACAGCCTCGATT 290
Db 278 ThrThrLeuLysSerAlaProThrThrProLysGluProAlaProThrThrLysSer 297
Qy 291 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
Db 298 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 317
Qy 321 CCGGACGCTGCTGTTCAACTCAATGACCTTGGCGGAGAACGCGGCTGCGGTACGCC 380
Db 318 ProLysGluProAlaProThrThrThrLysGluProAla----- 330
Qy 381 TGCTGTTGGAGCAACAACGCGGGTGGAGGATCTCCTGCTCGCCCTCGACAGCC 440
Db 331 -----ProThrThrLysSerAlaProThrThrProLysGluProAlaProThr 347
Qy 441 ATCCCAATATCAAGTGGCTGTCAACCCCTTCGCTACGCAAAATGGCGGCACTCG 500
Db 348 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 366
Qy 501 GCTACCTGACGACTTCCCGCGCTCAACCGCGCATGCAACAATACTTTACCG--- 557
Db 672 LysGluProAlaProThrThrProLys 680
```

```
Db 367 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 386
Qy 558 -----CCGACACCC----- 566
Db 387 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 406
Qy 567 ---CGCGCACCATCTCGGCGGACCAATATCGGCGGCAATATCTCAAACTCGGTGAGG 623
Db 407 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 426
Qy 624 ACACCGTTT-----TCGCGCAGCTGCACATCTCTGCCACCGCGAGCGTCTCGCGGAAG 677
Db 427 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 446
Qy 678 TATCGCAGCTTCGACCGCTACTGGGCAAGCCATTCGGCCCAACAAAGCCGCGCATCA 737
Db 447 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 461
Qy 738 TCCGACGCGCAACATCGGCAAGGCTCTCAAGCACTCGGATACAAACGACGAAACATCCA 797
Db 462 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 477
Qy 798 -----GACACGCGCTCTCGCTACCGCGAAACCGTCGAAACAGTGCCTCTTACC 848
Db 478 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 497
Qy 849 AAAAAATACAGACGGGAGCATCGACTGGCAGAGCGTCCAAACCGCTGATCAGCG--- 905
Db 498 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 514
Qy 906 -----ACACCCCTGCAAAAGACTCGACCGCGGACCGCGCAACACCGC 947
Db 515 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLys 532
Qy 948 CGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCGCAACGCGTCTATCTGTTT 1007
Db 533 LeuThrPro-----ThrThrProGluLysLeuAla----- 542
Qy 1008 CACCTATTTCGTCCTACAAATCCGGCACAGCAGCTGGCAAAACTGGTCGAGGACG 1067
Db 543 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 557
Qy 1068 GCATAGACGTTACGCTCTGACCAACTCGCTACAGCGGACGCGTTCGCG----- 1118
Db 558 AlaProThrThrProGluGluProThr-----ProThrThrProGluGluPro 573
Qy 1119 ---CGTTCATTCGGCTACGTCAATACGAAACCGCTGCTCAAGCGCGGATCAAC 1175
Db 574 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 590
Qy 1176 TCTACGAGCTGCAACCAACCATGCGTCCCGCCCAAAAGACAAAGGCTGACCGCA 1235
Db 591 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 603
Qy 1236 GCTCCGTAACGACGCTGATGCCAAACCTTATTGTGGACGCGCAACGATCTTCATCG 1295
Db 604 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 622
Qy 1296 GCTCATTCACTCGACCCCGCTTCGCGACGCTCAATACCGAAATGGGCGTCTCATCG 1355
Db 623 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 636
Qy 1356 AAGCCCAAAATCGCAGAACACATGGAGCGGACCTCGCGGATACACACCGCAATACG 1415
Db 637 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 652
Qy 1416 CTTACCGCTTACCTCGCAAAACCAACCGCTGCAATGGCAGCTCCCG---CCACCC 1472
Db 653 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 671
Qy 1473 GAAAAACCTTACCGAAGACCGCAAG 1499
Db 672 LysGluProAlaProThrThrProLys 680
```



```
QY 1176 TCTACGAGTGCACCCCAACCATGCGTCCCGCCGACAAAAGACAAAGAGCCCTGACCGGCA 1235
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
682 -----ProThrThrProLysGluProAlaProThr-----ProLys 694
QY 1236 GTCGTAACAGCCTGATGCCAAACCTTCATTTGTGGAGGCGAAGCATCTTCATCG 1295
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
695 GluProAlaProThr-----ThrProLysGluThrAlaProThrProLysGlyThrAla 713
QY 1296 GCTCATTTCAACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGCGTGTGTCATCG 1355
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
714 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 727
QY 1356 AAAGCCCCAAATCCGACAGATGGAGCGACCTCGCGATACCAACCCGAATACG 1415
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
728 ProAlaProLysGlu-----LeuAlaProThrThrLysGluProThrSer 743
QY 1416 CCTACCGGTTACCTCGACAAACACACACCGCTGCAATGCGACGATCCCG---CCACCC 1472
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
744 ThrThrSerAspLysProAlaProThrThrPro-----LysGlyThrAlaProThrPro 762
QY 1473 GAAAAACCTACCCGACAAACCCGGAAG 1499
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
763 LysGluProAlaProThrThrProLys 771

RESULT 13
US-11-169-232-44
; Sequence 44, Application US/11169232
; Publication No. US20060025570A1
;
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
;
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;
; NUMBER OF SEQUENCES: 143
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseir, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
;
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1270 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-11-169-232-44
Alignment Scores:
Pred. No.: 5.65e-09 Length: 1270
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-44 (1-1270)
QY 144 CCAAACTGCTCTCTCGAGCAACATCTGCAAAATCGGCGACACCCCTCATTAACAACGGGC 203
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
197 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 214
QY 204 TATCCGACATCTACC-----TGCTCGAGAGCCCGCCACGAGCCCTTGCCTG 248
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 234
QY 249 -----CCGCGCGCCCTTATCGAATCTGCCGACACACAGCCCTCGATT 290
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 254
QY 291 TGCATATCTACA-----TTTGGCGCAACGACATTT 320
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 274
QY 321 CCGCGAGCTGCTTCAACCTCATGTACCTTCCGCGAGAACGCGCGTGCCTGACGCC 380
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
275 ProLysGluProAlaProThrThrLysGluProAla----- 287
QY 381 TGCTGTTGGACGACACACACGCGGTTGGACGATCTCTCTCGCCCTCGACAGCC 440
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
288 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 304
QY 441 ATCCCAATATCGAGTGGCGCTGTTCAACCCCTTCTGCTACGCAAAATGCGCGCATCG 500
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 323
QY 501 GCTACCTGACCGACTCCCGCGCTCAACCGCGCATGCGCAACAAATCTTTACCG--- 557
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
324 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 343
QY 558 -----CCGACAAAC----- 566
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
344 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 363
QY 567 ---GCGCCACCATCTCGGCGGACGCAATATCGGCGACGAATCTTCAAAGTCGGTAGG 623
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 383
QY 624 ACACCGTT-----TCGCGACCTGACATCTCGCCACCGGCGAGCGTGTGCGCGAAG 677
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 403
QY 678 TATCGCAGCACTTCGACCGCTACTGGGCAAGCCATTCGCGCCACACGCGCGCATCA 737
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 418
QY 738 TCGCGACGGCAACATCGGCAAGGGTCTTCAAGACTCGGATCAACAGCAAGAAATCCA 797
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 434
QY 798 -----GACACGCGCTCTCGCTACCGGAAACCGTCGAACAGTCCCGCTCTACC 848
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 454
QY 849 AAAAAATACAGACGGGACGCGATCGACTGGCGAGCGTCCAAACCGCGCTGATCAGCG--- 905
```


Db 385 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 404
QY 567 ---GGCCACCATACTCGCGGAGCAATATCGGCGAGCAATCTTCAAAGTCGGTGAGG 623
Db 405 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 424
QY 624 ACACGGTTT-----TCGCCGACCTGGACATCTCTCGCCACCGGCGAGCGCTCGCGCGAAG 677
Db 425 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 444
QY 678 TATCGACGACTTCACCGCTACTCGGCAAGCATTCCGCCACCAACCGCGACGCGCATCA 737
Db 445 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 459
QY 738 TCCGCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACACGACGCAACGCAATCCA 797
Db 460 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 475
QY 798 -----GACACGCGTCTCGCTACCGCGAAACCGTCGAACAGTCGCCCTCTACC 848
Db 476 AlaProThrThrProLysProAlaProThrThrProLysGluProAlaProThrThr 495
QY 849 AAAAAATACACGGGCGCATCGACTGCGAGCGTCCAAACCGCGCTGATCAGCG--- 905
Db 496 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 512
QY 906 -----ACACCCCTGCAAAAGGACTCGACCGCGCGACCGCGCAACCGC 947
Db 513 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys 530
QY 948 CGATTGCGGGAGGCTCAAGACGCGCTCAACAGCGCGGAAAGCGCTCTATCTGGTTT 1007
Db 531 LeuThrPro-----ThrThrProGluLysLeuAla----- 540
QY 1008 CACCTATTTCGCTCAAAAATCCGGCAGACGCACTCGCAAACTGTGCGAGACG 1067
Db 541 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 555
QY 1068 GCATAGAGTTTACCGTCTGACCACTCGCTACAGGCGACGAGTTGCGG----- 1118
Db 556 AlaProThrThrProGluLuproThr-----ProThrThrProGluLupro 571
QY 1119 ---CCGTCAATTCGGGTACGTAATAATACCGAAACCGCTCTCAAAAGCGGCAATCAAC 1175
Db 572 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 588
QY 1176 TCTAGAGCTCAACCCACCATCGCTCGGCGGCGGCAAAAGAGCGCTGACGGCA 1235
Db 589 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 601
QY 1236 GCTCGTAACCGCTGCAATCCCAAAACCTTCATTGTGCGCGGCAACGCACTTCATCG 1295
Db 602 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 620
QY 1296 GCTCATTTCAACCTCGACCCCGTTCGCGACGCGCTCAATACCGAAATGGGCGTCTCATCG 1355
Db 621 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 634
QY 1356 AAAGCCCAAAATCGCAGACAGATGGAGCGCACCTCGCGGATACACACCCGGAATACG 1415
Db 635 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 650
QY 1416 CCTACCGGTTTACCTCGACAAACACACCGCTGCAATGCGAGATCCCG---CCACCC 1472
Db 651 ThrThrSerAspLysProAlaProThrThrPro-----LysGlyThrAlaProThrPro 669
QY 1473 GAAAAACCTTACCCGAACGCCGAAG 1499
Db 670 LysGluProAlaProThrThrProLys 678

RESULT 15

US-11-169-232-142

; Sequence 142, Application US/11169232

; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-11-169-232-142
Alignment Scores:
Pred. No.: 5,67e-09 Length: 1313
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-142 (1-1313)
QY 144 CCAAACTGCTCTCGGAGCAATCTGCAATCCGGCACACCCCTCATTAACACGGGC 203
Db 240 ProThrProLysAlaGluThrThrThr-----LysGlyProAlaLeuThrThrProLys 257
QY 204 TATCCGACATCTACC-----TGCTCGACGACCCCGACGAGCCCTTCGCG 248
Db 258 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 277
QY 249 -----CCCGCGCGCCCTTATCGAATCTGCGGAACACAGCGCTCGATT 290
Db 278 ThrThrLeuLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 297

```
Qy 291 TGCAATACTACA-----TTTGGCGCAACGACATT 320
Db 298 AlaProThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 317
Qy 321 CCGGAGGCTGCTGTTCAACCTCATGTACCTTGGCGCAGAACGGCGTGGCGTACGCC 380
Db 318 ProLysGluProAlaProThrThrThrLysGluProAla----- 330
Qy 381 TGCTGTTGGACGCAACAACACGCGCGGTGGACGATCTCTGCTCGCCCTCGACAGCC 440
Db 331 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 347
Qy 441 ATCCCAATATCGAAGTGGCGCTTCAACCGCTTCTGCTAGCGAATATGGCGGCGACTCG 500
Db 348 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 366
Qy 501 GCTACCTGACGACTTCCCGCGCTCAACCGCGCATGCACAACAAATCTTTACCG-- 557
Db 367 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 386
Qy 558 -----CGCACAAAC----- 566
Db 387 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 406
Qy 567 ---GGCCACCATACTCGGGGAGCCGCAATATCGGGGACGAGATCTTCAAAATCGGTGAGG 623
Db 407 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 426
Qy 624 ACACCGTTT-----TCGCCGACCTGGACATCTCGCCACCGGAGCGTGTGCGCGAAG 677
Db 427 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrThrLysGluPro 446
Qy 678 TATCGCAGGACTTCGACCGCTACTGGCAAGCATTCGCCGCCACAAACGCCACGCGCATCA 737
Db 447 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 461
Qy 738 TCCGAGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACAAACGACGAAACATCCA 797
Db 462 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 477
Qy 798 -----GACACGCGTCTCGCTACCGCGGAAACGTCGAACAGTCGCCCTCTACC 848
Db 478 AlaProThrThrProLysLysProAlaProThrThrThrProLysGluProAlaProThrThr 497
Qy 849 AAAAAATACAGACGGGCGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCG--- 905
Db 498 ProLysGluProAlaProThrThrThr-----LysLysProAlaProThrAlaPro 514
Qy 906 -----ACACCTTGCAAAAGGACTGACCGCGACCGCGCAACCGC 947
Db 515 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys 532
Qy 948 CGATTGCGGGAGGTGCAAGACGGCTCAAAACGCGGAAAGAGCTCTATCTGGTTT 1007
Db 533 LeuThrPro-----ThrThrProGluLysLeuAla----- 542
Qy 1008 CACCCTATTTCGTCCTCAAAATCCGGCAGACGCACTGGCAAAACCTGGTGCAGGACG 1067
Db 543 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 557
Qy 1068 GCATAGAGTTACCGTCTCGCAACTCGCTACAGCGGACCGAGCTTGGCG----- 1118
Db 558 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 573
Qy 1119 ---CCGTCCATTCCGGCTACGTCAAATACCGAAACCGGTGCTCAAGCGCGCATCAAC 1175
Db 574 AlaProThrThrProLysAlaAla---AlaProAlaProThrProLysGluProAla 590
Qy 1176 TCTACGAGTGCACCAACCAACCATGCGTCCCGCCCAAGAAAGCGCTGACCGGCA 1235
Db 591 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 603
```

```
Qy 1236 GCTCGTAACAGCAGCTGCATGCGCAAAACCTTCAATTGTGGAGCGCAACGCAATCTTCATCG 1295
Db 604 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 622
Qy 1296 GCTCATTCAACCTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGTCGTCATCG 1355
Db 623 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 636
Qy 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCAACCCGAATACG 1415
Db 637 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 652
Qy 1416 CCTACCGGTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCG---CCACCC 1472
Db 653 ThrThrSerAspLysProAlaProThrThrPro-----LysGlyThrAlaProThrThrPro 671
Qy 1473 GAAAAACCTACCCGAAACGACCCGAAG 1499
Db 672 LysGluProAlaProThrThrProLys 680

RESULT 16
US-11-169-232-50
; Sequence 50, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-Jan-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-Jun-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-Dec-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-Aug-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Ceert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
```

US-11-169-232-50

Alignment Scores:

Pred. No.: 5,67e-09 Length: 1314
 Score: 247.50 Matches: 135
 Percent Similarity: 35.2% Conservative: 44
 Best Local Similarity: 26.5% Mismatches: 205
 Query Match: 8.7% Indels: 125
 DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-50 (1-1314)

QY 144 CCAACCTGCTCTCTGACAAACATCTGCAAAATCCGCACACACCCCTCATACAAACGGCG 203
 DB 241 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 258
 QY 204 TATCGACATCTACC-----TGCTCGACGACCCCGCAGAACGCTTGGCG 248
 DB 259 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 278
 QY 249 -----CCGCGCGCGCTTATCGAATCTGCGAATCTGCGAACACACAGCTCGATT 290
 DB 279 ThrThrLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 298
 QY 291 TGCAATACATA-----TTGGCGCAACGACATTT 320
 DB 299 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 318
 QY 321 CGGCGAGCTGCTTCAACCTCATGTACTTGGCGGAGAGCGGGCGGTGCGGTACGCC 380
 DB 319 ProLysGluProAlaProThrThrThrLysGluProAla----- 331
 QY 381 TGCTGTTGGACGACAAACACGCGGGTTGGAGATCTCTGCTGCGCCCTCGACAGCC 440
 DB 332 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 348
 QY 441 ATCCCAATATCGAAGTGGCTGTCAACCCCTTGTCTACGCAAAATGGCGCGCACTCG 500
 DB 349 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 367
 QY 501 GCTACCTGACGACTTCCCGCTCAACCGCGCATGCAACAAATCCTTTACCG--- 557
 DB 368 GluProThrProThrProLysGluProAlaProThrThrLysGluProAlaProThr 387
 QY 558 -----CCGCAACAC----- 566
 DB 388 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 407
 QY 567 ---CGGCCACCATCTCGGCGAGCGCAATATCGGCGAGCAATACTTCAAAAGTCGGTGAGG 623
 DB 408 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 427
 QY 624 ACACGGTT-----TCGCGACCTGGACATCTCGGCACCGGAGCGTGTGCGGCGAAG 677
 DB 428 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 447
 QY 678 TATCGCAGACTTCGACCGCTACTGGGCAAGCCATTCGCGCCCAACGCGACGCGCATCA 737
 DB 448 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 462
 QY 738 TCCGACGGCAACATCGGCAAGGCTTCTCAAGCACTCGGATACAAACGAGCAAAACATCCA 797
 DB 463 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 478
 QY 798 -----GACACGGCTCTGCTGCTACCGCGAAACGGTGGAAACAGTCGCGCTCTTACC 848
 DB 479 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 498
 QY 849 AAAAAATACAGACGGGCGCATCTGACTGGCAGAGCGTCCAAACCGCTGATCAGCG--- 905
 DB 499 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlapro 515
 QY 906 -----ACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGC 947

Db 516 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLys 533
 QY 948 CGATTGCGGGAGGCTGACAGCGCTCAACAGCGCGCAAAAGCGCTCTATCTGGTTT 1007
 Db 534 LeuThrPro-----ThrThrProGluLysLeuAla----- 543
 QY 1008 CACCCTATTTCGCTCCCTACAAAATCCGCGACAGACGACTGGCAAACTGGTGCAGGACG 1067
 Db 544 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 558
 QY 1068 GCATGAGCGTTACCGTCTCTGACCAACTCGCTACAGGCGACGAGTGTGCGC----- 1118
 Db 559 AlaProThrThrProGluProThr-----ProThrThrProGluPro 574
 QY 1119 ---CCGTCCATTCGGGTACGTCAAATACCGAAACCGCTCTCAAGCCGCGCTCAAAAC 1175
 Db 575 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 591
 QY 1176 TCTACGAGCTGCAACCCCAACCATCGCTCCCGCCACAAAAGACAAAGGCTGACCGCA 1235
 Db 592 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 604
 QY 1236 GCTCCGTAACGAGCTGCTGCAAAACCTTTCATTGTGGCGCAACGCAATCTTCATCG 1295
 Db 605 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 623
 QY 1296 GCTCATTCAACCTGACCCCGTTCGCGAGCGCTCAATACGAAATGGGCGCTGCTCATCG 1355
 Db 624 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 637
 QY 1356 AAAGCCCAAAATCGCAGACAGATGAGCGCACCTCGCGATACACACCCGCAATACG 1415
 Db 638 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 653
 QY 1416 CCTACCGCTTACCTCGCAAAACCGCTTCAATGSCAGATCCCG---CCACCC 1472
 Db 654 ThrThrSerAspLysProAlaProThrPro---LysGlyThrAlaProThrThrPro 672
 QY 1473 GAAAAACCTACCGAAGCCCGAAG 1499
 Db 673 LysGluProAlaProThrThrProLys 681

RESULT 17
 US-11-169-232-46
 ; Sequence 46, Application US/11169232
 ; Publication No. US20060025570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Turner, Katherine
 ; Clark, Stephen C.
 ; Jacobs, Kenneth
 ; Hewick, Rodney M.
 ; Geshner, Thomas G.
 ; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
 ; NUMBER OF SEQUENCES: 143
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/11/169,232
 ; FILING DATE: 28-Jun-2005
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/124,557

1296	QY	GCTCATTTCAACCTCGAGCCCGTTCGCCAGCGCTCAATACGAAATGGGGCGTCGTCA	TCG	1355
630	Db	ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----	Lys	643
1356	QY	AAAGCCCCAAATCGCAGACAGATGGAGCGCACCTCGCGATACACACCCGGAATACG		1415
644	Db	ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer		659
1416	QY	CCTACCGCGTTACCTCGACAAACACACCGCCCTGCAATGGCAGATCCCG---	CCACCC	1472
660	Db	ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro		678
1473	QY	GAAAAACCTTACCGCAACCCGAAG		1499
679	Db	LvsGluProAlaProThrThrProLvs		687

```

; RESULT 19
; US-11-169-232-48
; Sequence 48, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
;

```

```

, ZIP: 021130
,
, COMPUTER READABLE FORM:
,
,     MEDIUM TYPE: Floppy disk
,
,     COMPUTER: IBM PC compatible
,
,     OPERATING SYSTEM: PC-DOS/MS-DOS
,
,     SOFTWARE: Patent In Release #1.0, Version #1.25
,
,
,     CURRENT APPLICATION DATA:
,
,         APPLICATION NUMBER: US/11/169,232
,
,         FILING DATE: 28-Jun-2005
,
,         CLASSIFICATION: <Unknown>
,
,
,     PRIOR APPLICATION DATA:
,
,         APPLICATION NUMBER: US/10/124,557
,
,         FILING DATE: 16-Apr-2002
,
,         APPLICATION NUMBER: US 07/643,502
,
,         FILING DATE: 18-JAN-1991
,
,         APPLICATION NUMBER: US 07/546,114
,
,         FILING DATE: 29-JUN-1990
,
,         APPLICATION NUMBER: US 07/457,196
,
,         FILING DATE: 29-DEC-1989
,
,         APPLICATION NUMBER: US 07/390,901
,
,         FILING DATE: 08-AUG-1989
,
,

```

```

,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: Csert, Luann
,
, REGISTRATION NUMBER: 31,822
,
, REFERENCE/DOCKET NUMBER: GI 5190
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: (617)876-1170
,
, TELEFAX: (617)876-5851
,
, INFORMATION FOR SEQ ID NO: 48:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 1354 amino acids
,
, TYPE: amino acid
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: protein
,
, SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-11-169-232-48

```

Alignment Scores:
Pred. No.: 5.69e-09 Length: 1354

Score:	247.50	Matches:	135
Percent Similarity:	35.2%	Conservative:	44
Best Local Similarity:	26.5%	Mismatches:	205
Query Match:	8.7%	Indels:	125
DB:	11	Gaps:	29
US-10-665-990A-13 (1-1561) x US-11-169-232-48 (1-1354)			
Qy	144	CCAAACGTGCTCTGGGACACATCCTCGAATCCGGCACACCCCTCTATACACAGGCG	203
		::: :::	
Db	281	ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys	298
Qy	204	TATCCGACATCTACC-----TGCTCGACGACCCCGACGAGCCCTTGCCG	248
Db	299	GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro	318
Qy	249	-----CCGCGCGCCCTTATCGAATCTGCCGACACAGCCTCGATT	290
Db	319	ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer	338
Qy	291	TGCAATACTACA-----TTTGGCGCAACGACATTT	320
Db	339	AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr	358
Qy	321	CCGGCAGGCTGCTTCAACCTCATCTTCCCGCGAAGACCGCGGTGCGGTACGCC	380
Db	359	ProLysGluProAlaProThrThrThrLysGluProAla-----	371
Qy	381	TGCTGTTGGAGCACACACACGCGGGTTGGACGATCTCTGCTGCGCCTCGACAGCC	440
Db	372	-----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr	388
Qy	441	ATCCCAATATCGAAGTGCCTCTTCAACCCCTTCGTCCTACGCAATGCGCGCACTCG	500
Db	389	ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys	407
Qy	501	GCTACCTGACCGACTTCCCGCGCTCAACGCGCGCATGCACACAAATCTTTTACCG---	557
Db	408	GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr	427
Qy	558	-----CCGCAACCC-----	566
Db	428	ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu	447
Qy	567	---GCGCCACCATCTCGCGGACGCAATATCGCGACGAATACTTCAAAGTCGGTAGG	623
		:::	
Db	448	ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr	467
Qy	624	ACACCGTTT-----TCGCGACCTGGACATCTCGCCACCGCAGCGTCTCGGCGAAG	677
		:::	
Db	468	ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro	487
Qy	678	TATCGACGACTTTCGACGCTACTGGGCAAGCATTCGCCCAACGCGCACCGCATCA	737
		:::	
Db	488	AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro	502
Qy	738	TCCGACGCGCACATCTCGGCAAGGCTTTCAGCATCTCGGATCAACAGCAACATCCA	797
		::: :::	
Db	503	SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro	518
Qy	798	-----GACACGGCTCTCTCGCTACCGCGAAACCGTCGAACAGTCCGCCCTCTACC	848
Db	519	AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr	538
Qy	849	AAAAAATACAGCGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCG---	905
		:::	
Db	539	ProLysGluProAlaProThrThrThr-----LysLysProAlaProThrAlaPro	555
Qy	906	-----ACACCCCTTCGAAAAGGACTTCGACCGCGACCGCGCGCAACCGC	947
Db	556	LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys	573
Qy	948	CGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCCGGAAAGACGCTATCTGCTTT	1007

```
Db 574 LeuThrPro-----ThrThrProGluLysLeuAla----- 583
QY 1008 CACCCTATTTCGTCCTTACAAATCCGGCACAGCGCACTGGCAAACTGTGTGAGGACG 1067
Db 584 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 598
QY 1068 GCATAGACGTTACCGTCTCGACCAACTCGCTACAGCGCAGCGAGCTTCCGCG----- 1118
Db 599 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 614
QY 1119 ---CGTCCATTCGGGTACGTCATAATACCGAAACCGCTGCTCAAGCCGCGCATCAAC 1175
Db 615 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 631
QY 1176 TCTAGAGCTCAACCCCAACCATCGCTCCCGCCGACCAAAAGAGCAAGGCGCTGACGGCA 1235
Db 632 -----ProThrThrProLysGluProAlaProThr-----ProLys 644
QY 1236 GCTCGTAACGAGCTGCATGCCAAACCTTCATTGTGGAGCGCAACGCAATCTTCATCG 1295
Db 645 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 663
QY 1296 GCTCATTCACCTGACCCCGTTCGCGACGCTCAATACCGAAATGGCGTGTGTCATCG 1355
Db 664 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 677
QY 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGATACACACCCGCAATACG 1415
Db 678 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 693
QY 1416 CCTACCGGTTACCTCGACAAACACACACCCGCTGCAATGGCAGATCCCG---CCACCC 1472
Db 694 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 712
QY 1473 GAAAAACCTACCCGAACGAAACCCGAAG 1499
Db 713 LysGluProAlaProThrThrProLys 721

RESULT 20
US-11-169-232-40
; Sequence 40, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
```

```
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-11-169-232-40

Alignment Scores:
Pred. No.: 5,7e-09 Length: 1361
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-40 (1-1361)
QY 144 CCAACCTGCTCTCGTGGACAAACATCTGCAATCGGCACACACCCCTCATACAAACGGC 203
Db 288 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 305
QY 204 TATCCGACATCTACC-----TGCTCGACGACCCCGCCAGAGCCCTTCCG 248
Db 306 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 325
QY 249 -----CCGCGCGCCCTTATCGAATCTGCGGAACACAGCCCTCGATT 290
Db 326 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrLysSer 345
QY 291 TGCAATACTACA-----TTTGCGCAACGACATTT 320
Db 346 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThr 365
QY 321 CCGGACGGTGTGTTCAACCTCATGTACCTTCCGCGAGAACGCGCGTGCAGTACGCC 380
Db 366 ProLysGluProAlaProThrThrLysGluProAla----- 378
QY 381 TGCTGTGGACGACAAACACACCGCGGGTGGACGATCTCTGCTCGCCTCGACAGCC 440
Db 379 -----ProThrThrLysSerAlaProThrThrProLysGluProAlaProThr 395
QY 441 ATCCCAATACGAAGTGGCGCTGTCAACCCCTTCGCTCTACGCAAAATGGCGCACTCG 500
Db 396 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 414
QY 501 GCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGACACAAATCTTTACCG--- 557
Db 415 GluProThrProThrProLysGluProAlaProThrThrLysGluProAlaProThr 434
QY 558 -----CCGCAACCC----- 566
Db 435 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 454
QY 567 ---GCGCACCATCTCGGCGGACGCAATATCGGCACGAATACTTCAAAGTCGGTGAGG 623
Db 455 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 474
QY 624 ACACCGTTT-----TCGCGACCTGGACATCTCGCCACCGGACGCGTCTCGCGGAAG 677
||||| :||||| ||| ||| :|||
```

```
Db 475 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 494
QY 678 TATCGACGACTTCGACCGCTACTGGCAAGCCATTCCGCCCAACCCACCGCGCATCA 737
Db 495 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 509
QY 738 TCCGAGCGGCAACATCGCAAGGTCTTCAAGCACTCGGATACAAACGACGAAACATCCA 797
Db 510 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 525
QY 798 -----GACACGCGCTCTGCGCTACCGGAAACGTCGAACAGTCGCGCCCTCTACC 848
Db 526 AlaProThrThrProLysProAlaProThrThrProLysGluProAlaProThrThr 545
QY 849 AAAAAATACAGCGGACGATCGACTGCGCAGAGCGTCCAAACCCGCTGATCAGCG--- 905
Db 546 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 562
QY 906 -----ACACCCCTGCNAAGGACTCGACCGCGACCGCGCCCAACCGC 947
Db 563 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrProLysLys 580
QY 948 CGATTCCCGGGAGGTGCAAGCGCGCTCAACACGCGCGAAAGCGTCTATCTGGTTT 1007
Db 581 LeuThrPro-----ThrThrProGluLysLeuAla----- 590
QY 1008 CACCTATTTCGCTCAAAATCCGCGACAGCGACTGCGCAAACTGGTGCAGGACG 1067
Db 591 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 605
QY 1068 GCATAGAGTTACGCTCGCAACTCGCTACAGCGACCGACGCTTGGCG----- 1118
Db 606 AlaProThrThrProGluGluPro-----ProThrThrProGluGluPro 621
QY 1119 ---CGTCCATTCCGCTACGTCAAATACCGAAACCGCTGCTCAAGCCGCGCATCAAC 1175
Db 622 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 638
QY 1176 TCTACGAGCTGCAACCAACCATGCGTCCCGCCCAACAAAGACGAGCGCTGACCGCA 1235
Db 639 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 651
QY 1236 GCTCGTAACGACGCTGATGCAAACTTCAATGTGACGCGCAACCGCATCTTCATCG 1295
Db 652 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 670
QY 1296 GCTCATTAACCTCGACCCCGCTCGCAGCGCTCAATACCGAAATGGCGTGCATCG 1355
Db 671 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 684
QY 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGGACCGCTCGCCGATACCACACCGAATACG 1415
Db 685 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 700
QY 1416 CCTACGCGCTTACCTCGACAAACACCGCGCTGCAATGGACGATCCCG---CCACCC 1472
Db 701 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 719
QY 1473 GAAAAACCTTACCCGACGAAACCGGAAG 1499
Db 720 LysGluProAlaProThrThrProLys 728
```

RESULT 21

```
US-11-169-232-52
; Sequence 52, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
```

```
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserti, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-11-169-232-52
Alignment Scores:
Pred. No.: 5,7e-09 Length: 1363
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-52 (1-1363)
QY 144 CCAAAACCTGCTCCTCGGACCAACATCCTGCAATCCGCGCACACCCCTCATACAACGGGC 203
Db 290 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 307
QY 204 TATCCGACATCTACC-----TCGTCGACGACCCCGAACGAGCCCTTGGCG 248
Db 308 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 327
QY 249 -----CCGCGCGCGCCCTTATCGAATCTGCGCAACACACGACCTCGATT 290
Db 328 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 347
QY 291 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
Db 348 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 367
QY 321 CCGGAGCGGTGCTGTCAACCTCATGTACCTTGGCGCAGAACGCGCGGTGCGCGTACGCC 380
Db 368 ProLysGluProAlaProThrThrThrLysGluProAla----- 380
```



```
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-2 (1-1404)
QY 144 CCAACCTGCTCTCGGACCAACATCCGCAACACCCCTCATACAAACGGGC 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 331 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 TATCGACATCTACC-----TGCTCGAGACCCCGCAAGCCCTTGGCG 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 249 -----CCGCGCGCGCCCTTATCGAATCTGCGCAACACAGCCTCGATT 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 TGCAATATCTACA-----TTTGGCGCAACGACATTT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 389 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 CCGCAGGCTGCTTCAACCTCATCTGTTGCGCGAGACGCGGTGCGGTACGCC 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 409 ProLysGluProAlaProThrThrThrLysGluProAla----- 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 381 TGCTGTTGGACGACACACACGCGGGTTGGAGCATCTCTGCTCGCCCTCGACGCC 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 441 ATCCCAATATCAAGTGGCGCTGTTCAACCCCTCTGCTACGCAAAATGGCGGCACTCG 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 439 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 501 GCTACTGACGACTTCCCGCCCTCAACCGCGCATGCGCAACAAATCTTTTACCG--- 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 458 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 558 -----CCGCAACCC----- 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 478 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 567 ---GGCCACCATACTCGCGGACGCAATATCGGCGAGCAATCTCAAAGTCGGTGAGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 498 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 624 ACACGGTT-----TCGCGGACCTGGACATCTCGCCACCGGAGCTCGTCGCGGAG 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 518 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 678 TATCGACGACTTGCACGCTACTGGCAAGCCATTCCGCCCAACACGCGCGCATCA 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 538 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 738 TCCGAGCGGCAACATCGGCAAGGTCTTCAAGCATCTGGATACAAACGACGAAACATCCA 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 553 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 798 -----GACACGGCTCTCGCTACCGGCAACCGTCAACAGTCGCCCTCTACC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 569 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 849 AAAAATAACAGCGGACGATCGACTGCGAGCGTCCAAACCCGCTGATCAGCG--- 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 589 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 906 -----ACACCCCTGCAAAAGACTCGACCGCGCGACCGCGCAACACCGC 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 606 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 948 CGATTGCCGGAGGCTGCAAGACGCGCTCAACACGCGCGCAAAAGCTCTATCTGGTTT 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 LeuThrPro-----ThrThrProGluLysLeuAla----- 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1008 CACCTTATTTCGTCCTACAAATCCGGCACAGCGACTGGCAAAACTGGTGCAGGACG 1067
```

```
Db 634 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1068 GCATAGACGTTACCGTCTGTGACCAACTCGTACAGGCGACCGAGCTTGCCG----- 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 649 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1119 ---CGGTCCATTCGGCTACGTCAAATACCAAAACCGCTGCTCAAAGCCGCGCATCAAC 1175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 665 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1176 TCTACGAGCTGCAACCCCAACCATCGCTCCCGCCCAAAAAGACAAAGGCTGACCGGCA 1235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 682 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1236 GCTCCGTAAACGAGCTGCATGCCAAAACCTTCATTGTGGAGCGCAAAACGATCTTCATCG 1295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 695 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1296 GCTCATTCAACTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCGTCTCATCG 1355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGCATACACACCGCAATACG 1415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 728 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1416 CCTACCGGTTCCTCGACAAACACACACCCCTCGCAATGGCAGATCCCG---CCACCC 1472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 744 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1473 GAAAACTCCCAACGACCGCAAG 1499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 763 LysGluProAlaProThrThrProLys 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 23
US-11-169-232-62
; Sequence 62, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
```

; FILING DATE: 08-AUG-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cserr, Luann
 ; REGISTRATION NUMBER: 31,822
 ; REFERENCE/DOCKET NUMBER: GI 5190
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)876-1170
 ; TELEFAX: (617)876-5851
 ; INFORMATION FOR SEQ ID NO: 62:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1404 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
 US-11-169-232-62

Alignment Scores:
 Pred. No.: 1404
 Score: 5,728-09 Length: 1404
 247.50 Matches: 135
 Percent Similarity: 35.2% Conservative: 44
 Best Local Similarity: 26.5% Mismatches: 205
 Query Match: 8.7% Indels: 125
 DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-62 (1-1404)

```

QY 144 CCAACCTGTCTCGACCAACATCTGCAAAATCCGGCACACCCCTCATACACGGGC 203
    |||||
Db 331 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrProLys 348
    |||||
QY 204 TATCCGACATCTACC-----TGCTCGACGACCCCGCCCAAGCCCTTGCCG 248
    |||||
Db 349 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 368
    |||||
QY 249 -----CCGCGCGCCCTTATCGAATCTGCGCAACACAGCCTCGATT 290
    |||||
Db 369 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrLysSer 388
    |||||
QY 291 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
    |||||
Db 389 AlaProThrThrProLysGluProAlaProThrThrLysGluProAlaProThrThr 408
    |||||
QY 321 CCGCGAGGCTGCTGTCAACCTCATGTACCTTGGCGCAGAGACGGCGGTGCGGTACGCC 380
    |||||
Db 409 ProLysGluProAlaProThrThrThrLysGluProAla----- 421
    |||||
QY 381 TGCTGTGGACGACACACACGCGCGGTGGAGAGTCTCTGCTGCGCCCTCGACAGCC 440
    |||||
Db 422 -----ProThrThrLysSerAlaProThrThrProLysGluProAlaProThr 438
    |||||
QY 441 ATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCGTCTACGCAAAATGGCGGCACTCG 500
    |||||
Db 439 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrProLys 457
    |||||
QY 501 GCTACCTGACGAGCTTCCCGCGCCTCAACCCCGCATGCAACAAATCTTTTACCG--- 557
    |||||
Db 458 GluProThrProThrProLysGluProAlaProThrThrLysGluProAlaProThr 477
    |||||
QY 558 -----CCGACAACC----- 566
    |||||
Db 478 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 497
    |||||
QY 567 ---GGCCACCATACTCGGCGGCAATATATCGGCGAGCAATACTTCAAGTCGGTGAGG 623
    |||||
Db 498 ProAlaProThrThrProLysGluProAlaProThrThrLysGluProSerProThr 517
    |||||
QY 624 ACACCGTTT-----TCGCCGACCTGGACATCTCCGCCACCGCGGCGGTGCTCGCGCAAG 677
    |||||
Db 518 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 537
    |||||
QY 678 TATCGCAGGACTTCAGCGGTACTGGGCAAGCCATTCGCGCCCAACAGCGCACGGCGATCA 737
    |||||
  
```

```

Db 538 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 552
QY 738 TCCGCGAGCGCAACATCGGCAAGGTCTTCAAGCACTCGGATACAAACGAGAAACATCCA 797
    |||||
Db 553 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 568
    |||||
QY 798 -----GACACGGCGCTCCTCGCTACCGCGAAACCGTCGAACAGTCGCCCTCTACC 848
    |||||
Db 569 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 588
    |||||
QY 849 AAAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCG--- 905
    |||||
Db 589 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 605
    |||||
QY 906 -----ACACCCCTGCAAAAGGACTCGACCGGACCGCGCAACACCGC 947
    |||||
Db 606 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys 623
    |||||
QY 948 CGATTGCGGAGGCTGCAAGACGCGCTCAACACAGCCGAAAAAGCGTCTATCTGGTTT 1007
    |||||
Db 624 LeuThrPro-----ThrThrProGluLysLeuAla----- 633
    |||||
QY 1008 CACCTATTTCGTCCTTACAAAATCCGGCACAGCACTGCGCAAACTGGTGCAGGACG 1067
    |||||
Db 634 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 648
    |||||
QY 1068 GCATAGAGTTTACCGTCTGACCAACTCGCTACAGGCGACCGAGTTCGCG----- 1118
    |||||
Db 649 AlaProThrThrProGluGluProThr-----ProThrThrProGluGluPro 664
    |||||
QY 1119 ---CCGTCCATTCCGGCTACGTCAAATACCGAAAAACGCTCTCAAGCCCGGCATCAAC 1175
    |||||
Db 665 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 681
    |||||
QY 1176 TCTACGAGCTGCAACCCCAACATGCGTCCCGCCACAAAGAGCAAGGCGCTGACCGCA 1235
    |||||
Db 682 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 694
    |||||
QY 1236 GCTCCGTAAACGAGCTGCATGCCAAACCTTCATTGTGGAGGCGCAACGATCTTCATCG 1295
    |||||
Db 695 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 713
    |||||
QY 1296 GCTCATTCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGCTCGTCATCG 1355
    |||||
Db 714 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 727
    |||||
QY 1356 AAGCCCAAAATCGCAAGACAGATGAGCGCACCTCGCGGATACCAACCCGAATACG 1415
    |||||
Db 728 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 743
    |||||
QY 1416 CCTACCGGTTACCTCGACAAACACACACCGCCCTGCAATGGCAGATCCCG---CCACCC 1472
    |||||
Db 744 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 762
    |||||
QY 1473 GAAAAACCTACCCGCAAGCAACCCCGAAG 1499
    |||||
Db 763 LysGluProAlaProThrThrProLys 771
    |||||
RESULT 24
US-10-467-657-5700
; Sequence 5700, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
  
```

```
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5700
; LENGTH: 430
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5700

Alignment Scores:
Pred. No.:          9e-09          Length:          430
Score:             243.50         Matches:         123
Percent Similarity: 33.7%         Conservative:    37
Best Local Similarity: 25.9%      Mismatches:     166
Query Match:       8.5%          Indels:         149
DB:                9            Gaps:           24

US-10-665-990A-13 (1-1561) x US-10-467-657-5700 (1-430)
QY 310 CAACGACATTTCCGCGAGCTGCTGTTCAACCTCATGTACCTTCCGCGAGAACGGCGGT 369
D 21 GluGlnArgIleArgAsnAlaValGlyGlySerIleProPheSerGlyThrAsnThr 40
QY 370 GCGCGTACCGCTGCTGTTGGACGCAACAACACGCGCGGTGGACGATCTCCTCTCGC 429
D 41 LysHisArgProSerLeuGlyLysProValLeuPheHisIleSerAlaSerAspGly--- 59
QY 430 CCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAA---CCCTTTCGTCT---ACG 483
D 60 -----IleArgTyrGluLeuLeuAlaThrGlnThrProPheArgProLeuAla 75
QY 484 CAATGGCGCGCACTCGGCTACTGACCGACTTCCCGCGCTCAACCGCGGATGACCAA 543
D 76 GlnLeuLeuArgThr-----LeuProProPro---ProProHisThrCys 89
QY 544 CAATCTCTTACCGCGCAACCG---CGCCACCTACT-----CGG 579
D 90 ArgAlaAlaArgArgAsnArgProValArgHisArgThrArgProAlaThrProProPro 109
QY 580 -----CGCGGAGCGCAATAT-----CGG 597
D 110 ThrArgArgMetAspArgAspArgLeuArgArgProArgHisAlaProValProArg 129
QY 598 CGACGAATCTTCAA-----ATCGGTGAGGACACCGTTTTCGCGCA--- 639
D 130 ArgAspLeuLeuGlnGlyGlyGlyThrTyrAlaAlaGlyTyrHisArgAlaGlyArgGly 149
QY 640 -----CCTGGACATCTCTCGC 654
D 150 PheGlyArgPheMetAlaGluProAlaLeuPheProArgGlnProProLeuLeuProAsp 169
QY 655 CACCGCGACGTCGTCGGCGAAGT-----ATCGCAGCATCTCGACCG 696
D 170 HisArgHisGlyLysArgThrGlyArgLeuGlyGlyArgGlnLysArgLeuArgPro 189
QY 697 CTACTGGGAAGCATTCGCGCCCAACACCCACGCGCATCATCCGAGCGGCAACATCGG 756
D 190 TyrAlaGlyGly-----AlaAspAspValHisAlaHisArgArgGlnArgGlnArgMet 207
QY 757 CAAGGCTCTTCAAGCACTCGGATACACGACGAAACATCCAGACAGCGCTCTCTCGGCTA 816
D 208 AlaArg-----GlnArgProAspAlaArgAspGluArgProHisArg 221
QY 817 CGCGGAAC-----CGTGAACAGTCGCGCTCTTACCAAAAAAATACAGCGGAGC 867
D 222 ArgArgHisArgHisCysArgArgGlnThrAlaAlaAlaGluIleHisThrAspValAla 241
QY 868 CATCGACTGGCAGAGGTCCTCAACCGCTGTATCAGCGACACCCCTGCAAAAGGACTCGA 927
D 242 Phe-----HisAlaCysArg-----Gln 247
QY 928 CCGCGACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGACGCGCTCAACAGCGCCGA 987
D 928 -----Gln 987
```

```
Db 248 ProGlyArgLeuGlnGlnAsnAspCysArgAsnGlnGlnArgGlnAlaTyrAspAlaArg 267
QY 988 AAAAAAGCGTCTATCTGTTTCAACCTATTTCCGTCCCTACAAAATCCGCGCACAGACGCACT 1047
D 267 ----- 267
QY 1048 GGCAAAACTGTCAGGACGCATAGAGTTACCGTCTGTGACCACTCGCTACAGGCGAC 1107
D 268 -----ThrPheGlyAlaGluTyrGlyGlnAsnAlaProAsnGlnArgThrHisGlyGln 285
QY 1108 CGACGTTGCCCGCTCCATTCGGCTACGTCACAAATACCGAAACCGCTGCTCAAAAGCCGG 1167
D 286 Lys-----ProGlnProProArg---ArgHisIleGlyArgLysProHisGlnProLeu 302
QY 1168 CATCAAACTCTACGAGCTGCAACCAACCATGCGCTCCCGCCCAAAAGACAAAGGCGCT 1227
D 303 HisAspGlySerHisAlaAlaArgProGlnAsnArgGlnHisArgAlaAlaPro 322
QY 1228 GACCGGAGCTCCGTAACGCTGCATGCGCAAAACCTTCATTGTGGAGCGCAACGCAT 1287
D 323 AspHis-----ArgArgGlnAlaAlaIleSer-----GlnThrGln 334
QY 1288 CTTCATCGGCTCATTTCAACCTCGACCCCGCTCCGCGACGGCTCAATACCGAAATGGCGCT 1347
D 335 ArgGlnArgAsnProAlaAlaArgProProLeuHisThrAlaProAsnArgProAla--- 353
QY 1348 CGTCATGAAAGCCCCAAAATTCGAGAACACATGGAGCGCACCCCTCGCGCATACACACC 1407
D 354 ThrAsnArgArgProHisGlnArgGlnThrArgProProHisProHisArgHisArgHis 373
QY 1408 CGAATACGSCCTA----- 1419
D 374 GlnProArgThrGlySerProArgArgThrProProLeuProMetAlaGlyLeuProLeu 393
QY 1420 -----CGCGCTTACCTTCGACAAACCAACCAACCG 1446
D 394 AlaGlnHisGlnTyrAlaSerGlyAsnPheArgProArgHisProAlaAlaThrHisPro 413
QY 1447 CCTGCAATGGCAGCA---TCCCGCGCACCCGAAACCACTACCGGAA 1488
D 414 ProGlnMetAlaGlyCysProArgThrProThrProAlaProLys 428

RESULT 25
US-10-467-657-7448
; Sequence 7448: Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7448
; LENGTH: 430
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7448

Alignment Scores:
Pred. No.:          9e-09          Length:          430
Score:             243.50         Matches:         123
Percent Similarity: 33.7%         Conservative:    37
Best Local Similarity: 25.9%      Mismatches:     166
Query Match:       8.5%          Indels:         149
DB:                9            Gaps:           24
```


US-10-665-990A-13 (1-1561) x US-10-467-657-7448 (1-430)

QY 310 CAACGACATTTCCGCGAGGCTGCTTCAACCTCATGTACCTTCCGCGAGAACCGGGCT 369
Db 21 GluGlnArgIleArgAsnAlaValGlySerIleProPheSerGlyThrAsnThr 40
QY 370 GCGGCTAGCGCTGCTGTGGACGACAAACACGCGGGTGGAGCATCTCCTGCTCGC 429
Db 41 LysHisArgProSerLeuGlyValLeuPheHisIleSerAlaSerAspGly--- 59
QY 430 CCTCGACAGCATCCCAATATCGAAGTGGCGCTGTCAA---CCCTTCGTCCT---ACG 483
Db 60 -----IleArgTyrGluLeuLeuAlaThrGlnThrProPheArgProLeuAla 75
QY 484 CAAATGGCGGCTACCTGCGCTACCTGACGACTTCCCGCTCAACGCGCATCGACAA 543
Db 76 GlnLeuLeuArgThr-----LeuProProPro-----ProProHisThrCys 89
QY 544 CAAATCCTTTTACCCTCGACACCG---CGCCACCATACT----- 579
Db 90 ArgAlaAlaArgArgAsnArgProValArgHisArgThrArgProAlaThrProProPro 109
QY 580 -----CGCGGAGCGCAATAT-----CGG 597
Db 110 ThrArgArgMetAspArgAspArgLeuArgArgProArgHisAlaProValProArg 129
QY 598 CGAGCAATACTCAA-----AGTCGTGAGGACACCGCTTTTCGCCGA--- 639
Db 130 ArgAspLeuGlnGlyGlyGlyThrTyrAlaAlaArgTyrGlyHisArgAlaGlyArgGly 149
QY 640 -----CCTGGACATCCTCGC 654
Db 150 PheGlyArgPheMetAlaGluProAlaLeuPheProArgGlnProProLeuProAsp 169
QY 655 CACCGGAGGCTGCTCGCGGAGT-----ATCGACCACTTCGACCG 696
Db 170 HisArgHisGlyLysArgThrGlyArgLeuGlyGlyArgGlnLysArgLeuArgPro 189
QY 697 CTACTGGCAAGCATTTCCGCCACACGCGCATCATCGGACGCGCAACATCGG 756
Db 190 TyrAlaGlyGly-----AlaAspAspValHisAlaHisArgArgGlnArgGlnMet 207
QY 757 CAAGGGTCTTCAAGCACTCGGATACAGACGAAACATCCAGACGCGCTCGCGCTA 816
Db 208 AlaArg-----GlnArgProAspAlaArgAspGluArgProHisArg 221
QY 817 CCGGAAAC-----CGTCGACAGTGCCTCTTACCAAAATACAGACGGGACG 867
Db 222 ArgArgHisArgHisCysArgGlnThrAlaAlaGluIleHisThrAspValAla 241
QY 868 CATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGGACACCCCTGCAAAAGCATCGA 927
Db 242 Phe-----HisAlaCysArg-----Gln 247
QY 928 CCGCGACCGCGCAACCGCGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCCGA 987
Db 248 ProGlyArgLeuGlnGlnAsnAspCysArgAsnGlnGlnArgGlnAlaTyrAspAlaArg 267
QY 988 AAAAAGCGTCTATCGTTTTCACCTATTTCGTCCTTACAAATTCGGCGACAGCGCACT 1047
Db 267 ----- 267
QY 1048 GGCAAACTGGTCAGGACGGCATAGAGTTACCGCTGACCACTCGCTACAGGGAC 1107
Db 268 -----ThrPheGlyAlaGluTyrGlyGlnAsnAlaProAsnGlnArgThrHisGlyGln 285
QY 1108 CGAGTTGCGCGCTCATTCGCGTACGTCAATACCGAAACCGCTGCTCAAAGCCGG 1167
Db 286 Lys-----ProGlnProProArg---ArgHisIleGlyArgLysProHisGlnProLeu 302
QY 1168 CATCAAACTCTAGAGTGAACCCCAACCATCGCTCCCGCCCAACAAAGACAAAGCCCT 1227
Db 1168 ----- 1227

Db 303 HisAspGlySerHisAlaAlaArgProGlnAsnArgGlnHisHisArgAlaAlaPro 322
QY 1228 GACCGGAGCTCGTAACACGCTGATGCCAAACCTTCAATTGTGGACGGCAACGCGAT 1287
Db 323 AspHis-----ArgArgGlnAlaAlaIleSer-----GlnThrGln 334
QY 1288 CTTTCATCGCTCATTTCAACCTCGACCCCTTCGCGACGGCTCAATACCGAAATGGGCGT 1347
Db 335 ArgGlnArgAsnProAlaAlaArgProProLeuHisThrAlaProAsnArgProAla--- 353
QY 1348 CGTCATCGAAAGCCCAAAATCGCAGAACAGATGAGCGCACCTCGCCGATACACACCC 1407
Db 354 ThrAsnArgArgProHisGlnArgGlnThrArgProProHisArgHisArgHis 373
QY 1408 CGAATACGCGCTA----- 1419
Db 374 GlnProArgThrGlySerProArgThrProProLeuProMetAlaGlyLeuProLeu 393
QY 1420 -----CCGGCTTACCCTCGCAACCAACCAACCG 1446
Db 394 AlaGlnHisGlnTyrAlaSerGlyAsnPheArgProArgHisProAlaAlaThrHisPro 413
QY 1447 CTTGCAATGCGACCA---TCCCGCCACCCGAAACCTACCCGAA 1488
Db 414 ProGlnMetAlaGlyCysProArgThrProThrProAlaProLys 428

RESULT 26

US-11-079-463-7133
; Sequence 7133, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 2005-03-14
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7133
; LENGTH: 503
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7133

Alignment Scores:
Pred. No.: 9,19e-09 Length: 503
Score: 243.50 Matches: 112
Percent Similarity: 37.0% Conservative: 69
Best Local Similarity: 22.9% Mismatches: 184
Query Match: 8.5% Indels: 125
DB: 11 Gaps: 16

US-10-665-990A-13 (1-1561) x US-11-079-463-7133 (1-503)

QY 32 TCCCGGATGAAACACGCGAGCTCATTTCCCTTTTATGCTCTCTCT- 78
Db 61 AsnProValLysThrMetAlaTyrIleLeuLeuMetPheLeuProValValGlyLeu 80
QY 79 CTGTTTCATCTTCTTCAATATCTCCCAACCTGCTCCCTCGGACAACTCTCTCA- 120
Db 81 ValPheTyrPhePhePhe-----GlyArgSerGlnArgArgGluLysIleIle 96
QY 121 GGAAGACCGCTCATTTCAATATCTCCCAACCTGCTCCCTCGGACAACTCTCTCA- 174
Db 97 GlyLysLysSerTyrAspArgLeuLeuLysLysProMetAlaGluTyrLeuAlaGlnAsn 116
QY 175 -----AATCCG 180
Db 117 CysCysGluThrProLysGluTyrGluArgLeuIleGlnLeuPheGlnAsnThrAsnGln 136

Qy	181	GCACACCCCTCAATAACAACGGGCTATCCGCACATCTACCTGCTCGACGACCCCCACGAAGC	240
Db	137	AlaPhe-ProPheGluGlyAsnArgValAspIleTyr-----ThrGlyGlyTyrSerLy	154
Qy	241	CTTTGGCCGCGCGCGCCCTTATCGAATCTCGCGGAACACACAGCTCGATTGTGCAATACTA	300
Db	154	sLeuGlnAlaLeuLeuArgGluLeuGlnLysAlaArgLeuHisIleHisMetGluTyr	174
Qy	301	CATTGTGGCGCAACGACATTTTCGGCAGGCTGCTGTTCACCTCATGTACCTTCGCCGAGA	360
Db	174	rIlePheGluAspAspProValGlyArgLeuValArgAspValLeuIleGluLysAlaAr	194
Qy	361	ACGGGGCTGGGTAGCCTGCTGTGTGGACGACACACACCGCGGGGTGACGATCT	420
Db	194	gGluGlyValGluValArgValIleTyrAspAspValGlyCysTrpHisValProHisAr	214
Qy	421	CCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCACACCCCTTCGTCTCT	480
Db	214	gPhePheGluGluMetArgAspAlaGlyIleGluValArgSerPhe-----	229
Qy	481	ACGCAATGGCGCGCACTCGGCTACCTGACCGCACTTCCCC-----GCCTCAA	528
Db	230	-----LeuLysValArgPheProLeuPheThrSerLysValAs	242
Qy	529	CCGCGCATGCACACAAATCCTTTACCGCGGACACCGCGCCACACATCTCGCGGACG	588
Db	242	nTyrArgAsnHisArgLysIleValIleAspGlyArgIleGlyPheIleGlyGlyMe	262
Qy	589	CAATATCGGCGCAAGTAATCTCAAAGTCGGTAGGACACCGCTTTTCGCCGACCTGGACAT	648
Db	262	tAsnLeuAlaGluArgTyrMetArgGlyPheSerTrpGlyIleTrpArgAspThrHisIl	282
Qy	649	CCTCGCCACCGGCAAGCTCTCGCGCAAGTATCGCACACTTC-----GACCGCTACTG	702
Db	282	eLeuLeuGluGlyLysAlaValHisGlyLeuGlnThrAlaPheLeuLeuAspTrpTyrPh	302
Qy	703	GGCAGCCATTCCGCCACACAGCCGCGCGCATCTCGCAGCGGCACATCGCGCAAGG	762
Db	302	eValAspArgThrIleuIleThrAlaSerArgTyrPheProLys-----	316
Qy	763	TCATTCAAGCACTCGGATACAAACAGCAAAACATCCAGACACGCTCTCGCTCGCTCGCA	822
Db	317	IleGluAlaTyrGly-----AsnSerLeuVal-----G	326
Qy	823	AACCGTCGAACAGTGCCTC-----TACAAAAAATACAGACGGGACGCATCGACTG	876
Db	326	nIleValThr-SerGluProIleGlyProTrpLysGluIleMetGlnGly-----	342
Qy	877	GCAGAGCTCCAAACCCGCTGATCAGCGACACCCCTGCAAAGGACTCGACCCGCGACCG	936
Db	342	-----	342
Qy	937	CCGCAACCGCGATTGCGGGAGCTGCAAGACGCGCTCAAAACAGCCCGCAAAAAAGCGT	996
Db	343	-----LeuThrValAlaIleSerGlyAlaLysLysTyrPh	354
Qy	997	CTATCTGGTTTCAACCTATTTCCTCAAAAATCCGGCACAGACGCATCGCAAAACT	1056
Db	354	eTyrMetGlnThrProTyrPheLeuProThrGluGlnIleLeuGlyAlaMetGlnThrAl	374
Qy	1057	GGTGCAGGACGGCATAGACGTTACCGTCTCGACCAACTCCGCTACAGCGACCGACGTTGC	1116
Db	374	aAlaLeuAlaGlyValAspIleArgLeuMetLeuPro-----GluHisAlaAspAsnAr	392
Qy	1117	CGCCGCTCCATTTCGGGCTACGTCAAATACCGAAAAACCGCTGCTCAAAAGCCGGCATCAA	1176
Db	392	gValThrHisLeuGlySerCysSerTyrLeuAlaAspValLeuArgAlaGlyValLysVa	412
Qy	1177	CTACGAGCTGCAACCCCAACCATGCGCTCCCGGCCACAAAGACAAAGGCTGACGGCAG	1236
Db	412	lTyrPheTyrTys-----LysGlyPhe-----	419

Qy	1237	CTCCGTAA	CAGCGCTGCATGCCAAACCTTCA	TGTGGACGCGCAACAGCATCTT	CATCGG	1239
Db	420	-----	-----	LeuHisSerTysLeuMetValSerAspMetLeuSerThrValG	435	
Qy	1297	CTCATTTCA	ACCTCGACCCCGGTTCCCGACGGCTCAATACCGAAATGGGCGTGGTCA	TGCA	1356	
Db	435	ySerThrAsnLeuAspPheArgSerPheGluHisAsn	PheGluValAsnAlaPheMet	455		
Qy	1357	AAGCCCCAA	AAATCGCAGACATG	1381		
Db	455	rAspMetGluThrAlaLeuGlnMet	463			
RESULT 27						
US-10-873-528-109						
; Sequence 109, Application US/10873528						
; Publication No. US20050276814A1						
; GENERAL INFORMATION:						
; APPLICANT: Microbial Technics Limited						
; APPLICANT: Gilbert, Christophe FG						
; APPLICANT: Hansbro, Philip M						
; TITLE OF INVENTION: Proteins						
; FILE REFERENCE: PWC/P21129WO						
; CURRENT APPLICATION NUMBER: US/10/873,528						
; PRIORITY FILING DATE: 2004-06-23						
; PRIOR APPLICATION NUMBER: US/09/769,787						
; PRIOR FILING DATE: 2001-01-26						
; PRIOR APPLICATION NUMBER: GB 9816337.1						
; PRIOR FILING DATE: 1998-03-27						
; PRIOR APPLICATION NUMBER: US 60/125164						
; PRIOR FILING DATE: 1999-03-19						
; NUMBER OF SEQ ID NOS: 388						
; SOFTWARE: PatentIn Ver. 2.1						
; SEQ ID NO 109						
; LENGTH: 1236						
; TYPE: PRT						
; ORGANISM: Streptococcus pneumoniae						
US-10-873-528-109						
Alignment Scores:						
Pred. No.:		1,04e-08		Length:		1236
Score:		243.50		Matches:		124
Percent Similarity:		42.7%		Conservative:		94
Best Local Similarity:		24.3%		Mismatches:		219
Query Match:		8.5%		Indels:		74
DB:		9		Gaps:		16
US-10-665-990A-13 (1-1561) x US-10-873-528-109 (1-1236)						
Qy	126	GCGTCA	TTCATAC	TCTTCCAACTGCTCTCTGGACACATCTCTGCAATCCGGCACA	185	
Db	551	AlaSerAlaSerIleSerAlaSerGluSerAlaSerThrSerAlaSerGluSerAlaSerThrSerAlaSer	570			
Qy	186	CCCTCATACA	CAACGGGCTATCGCATCTT-----	ACCTGCTCGACGACCCCAACG	236	
Db	571	ThrSerThrSerAlaSerAlaSerThrSerAlaSerGluSerAlaSerThrSerAlaSer	590			
Qy	237	AAGCCCTT	CGCGCGCGCCCTTATCGAATCTGCCGAAACACACCTCGATTCGAT	296		
Db	591	AlaSerAlaSerThrSerAlaSerAlaSerThrSerAlaSerThrSerAlaSerAlaSerAlaSer	610			
Qy	297	ACTACAT	TTTGGCGACACATTTCCGGCAGGCTGCTGTTCACCTCATGTACTGTCGCG	356		
Db	611	ThrSerAlaSerAlaSerThrSerAlaSerGluSerAlaSerThrSerAlaSerAlaSerAlaSer	630			
Qy	357	CAGAACG	CGCGTCCGCTAGTCTGTGTGGACGACAAACACACGCGCGGTTCGACG	416		
Db	631	AlaSerThrSerAlaSerAlaSerAlaSerThrSerAlaSerAlaSerAlaSerAlaSerThrSer	650			
Qy	417	ATCTCTG	TCTGCGCTCGACAGCATCCCAATATCG---AAGTGCCTGTTCAACCCCT	473		
Db	651	AlaSerAlaSerAlaSerThrSerAlaSerValSerAlaSerThrSerAlaSerAlaSer	670			
Qy	474	TCGTCTCA	CAAAATGGCGCGCACTCGGTACTGACCGCATTTCCCGCGCTCAACCGCC	533		

Db	671	AlaSerThrSerAlaSerAlaSerAlaSerAlaSerThrSerAlaSerGluSerAlaSerThrSer	690
Qy	534	GCATGTCACAAACAAATCCTTTACCGCGCGAACACCGCGCCACCATACTCCGGCGACGCAATA	593
Db	691	AlaSerAlaSerAlaSerThrSerAlaSerAlaSerAlaSerThrSerAlaSerAlaSer	710
Qy	594	TCGGCGGACGAATACTTCAAAGTCGGTGAGGACACCGTTT-----	632
Db	711	AlaSerThrSerAlaSerGluSerAlaSerThrSerAlaSerAlaSerAlaSerThrSer	730
Qy	633	-----TCGCCGACCTGGACATCCTCGCCACCGCGCAGCGTCGTCGGCGAGTATCGC	683
Db	731	AlaSerGluSerAlaSerThrSerAlaSerAlaSerAlaSerThrSerAlaSerAlaSer	750
Qy	684	ACGACTTCGACCGCTACTGGCGCAAGCCATTCGCGCCACACCGCCACGCGCATCTCCGCCA	743
Db	751	AlaSerThrSerAlaSerGlySerAlaSerThrSerThrSerAlaSerAlaSerThrSer	770
Qy	744	CGCGCAACATCGGCAGGGTC-----TTCAAGCACTCGGATACAAACGACGAAACAT	794
Db	771	AlaSerAlaSerAlaSerThrSerAlaSerAlaSerAlaSerIleSerAlaSerGluSer	790
Qy	795	CCAGACACGCGCTCTCGTGGCTACCGCGAAACCGTCGAACACTGCCCTCTACCAAAAAA	854
Db	791	AlaSerThrSerAlaSerGluSerAlaSerThrSerThrSerAlaSerAlaSerThrSer	810
Qy	855	TACAGACGGGACGATCGACTGGCAGACGCTCCAAACCGCGCTGATCAGCGACACCCCTG	914
Db	811	AlaSerGluSerAlaSerThrSerAlaSerAlaSerAlaSerThrSerAlaSerAla	829
Qy	915	CAAAAGACTCGACGGGACCGCGCGAAACCGCGATTCGCGGAGGCTGCAAGACGCGC	974
Db	830	-----SerAlaSerThrSerAla--SerAlaSerAlaArgGln--ValA;gArgPr	845
Qy	975	TCAAACAGCCCGAAAAAGCGTCTATCTGGTTTCAACCTATTTTCGTCCCTACAAATCCG	1034
Db	845	oGlnProValHis-----LeuAsnArgHisGlnProValAr	857
Qy	1035	GCACAGACGCACTGGCAAAACTGGTCAGGACGCGATAGAGTTTACCGTCCTGCACCACT	1094
Db	857	gGlnProGlnValLeuValHisGlnLeuGlnHisGlnArgValHisArgLeuGlnHi	877
Qy	1095	CGCTACAGGCGACGCGATTGCGCGCTGCCATTCGG--CTACGTCAAATACCGAAAC	1151
Db	877	sGlnProValProArgLeuGlnArgGlnProValArgGlnLeuGlnValPro-----	895
Qy	1152	CGTGTCTAAAGCGCGCATCAAACTCTACGAGCTGCACCCCAACCATGCCGTCCCGCGCA	1211
Db	896	-ValLeuGlnSerGlnHisGln-----GlnValLeuGlnProGlnHi	909
Qy	1212	CAAAAGACAAAGGCGCTACCGCGGAGCTCCGTAACAGCGCTGCATGTCAAACCTTCATTG	1271
Db	909	s--ArgGlnValPro-----ArgLeuGlnGlnAlaHisGlnHisLeuAsnGl	924
Qy	1272	TGGACGCGCAACCGCATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGCGACGGTCA	1331
Db	924	NAArgGlnAlaProGlnLeuGlnValProValArgGlnProGlnArg-----	941
Qy	1332	ATACCGAAATGGGGCTGTCATCAAAAGCCCCCAAAATCGCGAAGACATGGAGCGCACCC	1391
Db	942	-----ArgGlnValArgGlnProGlnGlnValLeuValHisGlnLeuGlnHi	957
Qy	1392	TCGGCGGATACCAACCCGAATACGCTTACCG-----	1422
Db	957	sGlnArgValHisArgLeuArgArgGlnProValHisGlnSerGlnValProValAr	977
Qy	1423	-----CGTTACCTCGACAAACACAAACCGCGCTGCA-----ATGGCAGC	1460
Db	977	gGlnLeuProHisGlnGlnValProArgLeuGlnGlnAlaProValArgArgLeuGlnGl	997
Qy	1461	ATCCCGCCACCGAAAAACCTACCCGAAACGAAACCCGA-----AGCCAAAC	1505

[illegible]

```
QY 556 CGCCGACACCCGCCCATATCTCGCGGACGCAATATCGCGCAGCAATATCTCAAAGT 615
   ::::|
Db 245 GinArgAlaGlyArgHisArgArgGlyGlyGlyArgArgGlyValGlnAsp 264
   ::::|
QY 616 CGGTGAGGACACCGTTTCGC-----CGACCT---GGACATCTCTGCCACCGG 660
   |||
Db 265 Arg---ArgHisArgGlyArgLeuGlnAlaGlnProGlnGlnGlyProArgHisArg 283
   |||
QY 661 CAGCGTCGT-----CGCGAAGTATCGCAGCACTCTCGACCG 696
   |||
Db 284 GlnArgHisValArgGlyLeuArgAlaArgGlyHisArgAspValArgArgGlnArgPro 303
   |||
QY 697 -----CTACTGGCAAGCCATCTCCG----- 717
   |||
Db 304 GlyArgProValGlyGlyProValArgLeuLeuArgGlyHisGluArgGlnAlaArg 323
   |||
QY 718 -----CCACAA----- 723
   |||
Db 324 ValHisGlyProProAspProGlnAlaGluAlaProProGlyValHisArgGlyArgGly 343
   |||
QY 724 CGCCACGCGCATCTCCG-----CAGCGGCAACATCGGCAAGGTCTTCAAGCACTCGG 777
   |||
Db 344 HisHisGlyAlaHisProGlyTrpGlnLeuLeuHisGluAlaGlyGlnGlyGluArg 363
   |||
QY 778 -----ATCAACGACGCAACATCCAGACACGCGCTCTCGCTACCGCAAAC 825
   ::::|
Db 364 AlaGlyProAlaAlaGluAlaGlnAlaGlyGlnValArgAlaPro----- 378
   |||
QY 826 CGTCGAACAGTCGCCCTCTACCAAAATACAGACGGGACGCATCGACTGCCAGAGCGT 885
   ::::|
Db 379 -----HisValAlaAlaVal 383
   |||
QY 886 CCAAAACCGCTGATCAGCGACACCCC----- 912
   |||
Db 384 AlaGlyProProAspArgGlyHisProArgArgHisGlnValHisArgAlaArgGlyGln 403
   |||
QY 913 TGCAAAAGACTCGACCG-----CGACCGCGCAACCGCCGATTCGCCGGAGGCT 963
   |||
Db 404 LeuArgGluArgGlnProGlyHisArgArgProProArg----- 416
   |||
QY 964 GCAACAGCGCTCAACAGCCCGCAAAAGCGTCTATCTGTTTCAACCTATTTCTGCTCC 1023
   |||
Db 417 -----GlnGlyAlaAlaArgArgGlnLeuProGlyHisPro----- 428
   |||
QY 1024 TACAAAATCCGCACAGACGCACTGGCAAAACTGGTGCAGGACGCATAGACGTTACCGT 1083
   |||
Db 429 -----HisArgArgValHisGlyGlnArgProProArgHisArgGlnHisArg 444
   |||
QY 1084 CTGTACCAACTCGCTACAGGACCGCAGCTTCGCCCGCTCCATTCGGCTACGTCAAATA 1143
   ::::|
Db 445 GlnAlaHisValArgAlaValLeuArgAlaArgGlnArgValLeuGlnGlnArgAlaHis 464
   |||
QY 1144 CGCAAAACCGCTGCTCAAGCCGGCATCNA----- 1173
   |||
Db 465 LeuGlnProGlyArgGlnProGlnProGlnProGlyLeuArgLeuGlnGlnHisArgAsp 484
   |||
QY 1174 -----ACTCTACGAGTCGAACCCCAACCATCGCTCCGCCGCCACAAA----- 1215
   |||
Db 485 ArgHisGlyLeuLeuLeuArgAlaProValProGlyGlnProHisGlnProArg 504
   |||
QY 1216 -----AGCAAAAGCTGACCGGACCGTTCGCCCGCTCCATTCGGCTACGTCAAATA 1263
   ::::|
Db 505 AlaGluArgGlyArgAlaGlnProGlyArgGluLeuProGlyProArgLeuGlyGlnGlu 524
   |||
QY 1264 CTTTCATTGGACGGCAA-----ACGCATCTTCATCGGCTCATTCAA 1305
   |||
Db 525 AspArgArgGlyAspArgHisProGluAlaHisValValHisLeuHisArgGlyAlaVal 544
   |||
QY 1306 CTT----- 1308
   |||
Db 545 ProGlyArgGlyProAlaProProArgGlyGluHisGlnGlyValGlyGluGluHisArg 564
   |||
```

```
QY 1309 -----CGACCC 1314
   |||
Db 565 AspProGlyGlyGlnGluGlyAlaAspHisGluProLeuGlyArgAlaLeuGlnArgPro 584
   |||
QY 1315 -----CGTTCCGACCGCTCAATACCGAAAT 1341
   |||
Db 585 LeuGlnArgGluGlyAlaAspGlnArgHisArgProArgGlyArgValHisValArgGly 604
   |||
QY 1342 GGGCGTCGTATCGAAAGCCC-----CAAAATCGCAGAACACATCGAGCG 1386
   |||
Db 605 GlyArgGlyGlnArgGlnProAlaAlaAspAlaGluAlaAlaArgArgAlaGlyGlyPro 624
   |||
QY 1387 CACCTCGCCGATACACACCGCAATACGCTACGCGT----- 1425
   |||
Db 625 ArgProGlnGlnArgArgAlaArgSerGlySerProProCysSerProArgSerProGly 644
   |||
QY 1426 -----TACCTCGCAAAACACCAACCGCTTCAATGCGCAGA 1461
   |||
Db 645 SerArgArgSerSerAlaArgCysCysProArgArgTrpArgProAlaTrpArgArg 664
   |||
QY 1462 TCCCGCCAC---CCGAAAAACCTACCCGAAACGA 1491
   |||
Db 665 ArgGlyHisArgProArgGlyGluProAspArg 675
   |||

RESULT 29
US-11-096-568A-23618
; Sequence 23618, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23618
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(375)
; OTHER INFORMATION: Ceres Seq. ID no. 12414168
US-11-096-568A-23618

Alignment Scores:
Pred. No.: 3e-08 Length: 375
Score: 235.50 Matches: 119
Percent Similarity: 35.2% Conservative: 46
Best Local Similarity: 25.4% Mismatches: 149
Query Match: 8.3% Indels: 156
DB: 11 Gaps: 20

US-10-665-990A-13 (1-1561) x US-11-096-568A-23618 (1-375)
QY 157 CTGTGACAAATCTCTGCAATCCGSCACACCCCTCATAAACACGGGCTATCGCATCTA 216
   ::::|
Db 4 ProAlaArgArgGlnHisSerProThrHis-----GlyValArgSerArg 18
   |||
QY 217 CTGTCTGACACCCCGCCGAGGCGCTTGCAGCGCCGCGCCCTTATCGATCTGCCGA 276
   |||
Db 19 AlaAlaGlyArg-----ArgGlyArgValProGlyArgAlaValArg 32
   |||
QY 277 ACACAGCCTCGATTGTGCAATACTACATTTGGCGCAACGACATTTCCGCGAGGCTGCTGT 336
   |||
Db 33 ValGlnArgArg----- 36
   |||
QY 337 CAACCTCATGTACTCTTCCGCGAAGACCGCGCGTGGCGGTACGCTGCTGTGGACGACAA 396
   |||
Db 37 -----HisGlyGlyAlaAlaArgAlaAlaGlyValArgAspAlaGlyAlaGlyArg 54
   |||
QY 397 CAACA-----CGCGCGGGTTGG---ACGATCTCC 422
   |||
```

```
Db 55 HisGlyArgLeuGluGlyArgAspValArgArgGlyProAlaGlyTipArgSerAlaSer 74
QY 423 TGCTCGCCTCAGACGATCCCAATATCGAAGTCGGCTGTTCAACCCCTTCGTCCTAC 482
Db 75 ThrSerProGlyThrAlaGlyProArgProGlyGlyAlaCysArgSer---SerSerThr 93
QY 483 GCAATGGCGGCACCTCGCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGCACA 542
Db 94 ThrThrAlaAlaAlaSerGlyPro---AlaProGlyProThrAlaArgThrThr 112
QY 543 ACAATCTTTACCGCGGACAAACCGGCACCATATCGCGGACGCAATATCGCGCAGC 602
Db 113 AlaCysAlaTtpProProThrTipAlaArgSerTtpAlaG--- 126
QY 603 AATACTTCAAAGTCGGTGAGACACCGTTCGCGGACCTGGACATCTCGCCACCGCGCA 662
Db 127 -----ArgThrThrAlaSerArgProSerThrAlaSerProProPro 140
QY 663 GCGTGTGCGGGAATATCGACGACTTCGACCGCTACTGGGCAAGCCATTCCGCGCCACA 722
Db 141 SerThrThrAlaArgArgSerCysGlyTtpArgGlyArgAlaAlaGlyThrProGly 160
QY 723 ACGCACCGCATATCGCGAGCGGCAACATCGGCAAGGCTTCAAGCACTCGGATACA 782
Db 161 SerProArgProProThrSerAlaAlaSerSerCysArgAla--- 174
QY 783 ACGAGAAACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCGAACAGTCGCCCC 842
Db 174 ----- 174
QY 843 TCTACAAAAAATACAGACGGGACGATCGACTGCGAGCGTCCAAACCCGCGTGATCA 902
Db 175 -----ThrArgProAlaAlaProSerArgThrThrSerProCys----- 187
QY 903 GCGACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCCCGCATG-----CCG 956
Db 188 -----AlaSerValAlaArgProArgThrTtpProPro 198
QY 957 GGAGGCTGCAAGACGCGCTCAACAGCCCGAAAGAGGCTATCTGTTTCAACCTATT 1016
Db 199 SerProCysGlyAspThrSerSerSerCysHisSerSerAlaAlaTtp----- 214
QY 1017 TCGTCCCTTACAAAATCCGGCACAGACGACTGGCAAAACTGCTGAGGACGGCATAGC 1076
Db 214 ----- 214
QY 1077 TTACCGTCTGACCAACTCGCTACAGCGACCGAGTTCGCGCGTCCATTCCGCTACG 1136
Db 215 -----ArgAlaArgGlyArgArgProSerAlaProProThrLeuSerSerThr 230
QY 1137 TCAATACCGAAACCGCTGCTCAAGCGCGCATCAACTCTACGAGTGC----- 1187
Db 231 GlyPro-----SerThrThrAlaThrGlyAlaCysArgCysArgAla 245
QY 1188 ---AACCCAAACATCGCTCCCGCCCAACAAAGACAAAGGCTGACCGGCA----- 1235
Db 246 ProArgProThrThrTtpPro-----ThrArgSerAlaProAlaArgArgArg 262
QY 1236 ---GCTCGGTAAACAGCTGCATCCCAAACTTCATTGTGGAGCGGCAAAAGCATCTTCA 1292
Db 263 TtpThrProSerSerArgProProTtpSerSerSerAlaAlaAlaThrSerSerThr 282
QY 1293 TCGGCTCATTAACCTGACCCCC-----GTTCCGACGCTCAATACCGAATGGCG 1346
Db 283 ThrAlaProSerThrThrProProGlySerGlyProGlySer---ProSerTtpSer 301
QY 1347 TCGTCATCGAAAGCCCAAAATCCAGAAAGATGAGGCGCACCTCGCGATACACAC 1406
Db 302 GlyThrSerThrAlaSerSerThrAlaSerSerProSerThrProGlyProThrProPro 321
QY 1407 CCG-----AATACGCTTACCGCGTTA---CCCTCGACA 1436
|||
```

```
Db 322 ProSerSerCysAlaSerSerSerAlaSerSerThrProThrAlaAlaSerThrSerThr 341
QY 1437 AACACACCGCCTGCAATGGCAGCATC 1463
Db 342 AspThrSerValSerAsnSerSerIle 350
RESULT 30
US-10-517-696-111
; Sequence 111, Application US/10517696
; Publication No. US20060051759A1
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto A.
; APPLICANT: Turner, Leah R.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and
; FILE REFERENCE: DEX-0432
; CURRENT APPLICATION NUMBER: US/10/517,696
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: US 60/389,327
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 171
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 111
; LENGTH: 1256
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-517-696-111
Alignment Scores:
Pred. No.: 3.81e-08 Length: 1256
Score: 235.00 Matches: 146
Percent Similarity: 36.0% Conservative: 55
Best Local Similarity: 26.2% Mismatches: 238
Query Match: 8.2% Indels: 119
DB: Gaps: 24
US-10-665-990A-13 (1-1561) x US-10-517-696-111 (1-1256)
QY 93 CATGGTTGCCCCCACTCGAAGACGAGCGGAAAGCCGTCATTTCAATATTCCTCAAAACCTG 152
Db 156 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 175
QY 153 TCCTCTGGACACATCTCGCAATCCGGCACCCCTCATACAAACGGGTATCCGACA 212
Db 176 HisGlyValThrSerAlaProAspThrArgProProGlySerThrAlaProProAla 195
QY 213 TCTACCTGCTCGACGCCCC---ACGAAGCCC-----TTGCCCGCCGCG 254
Db 196 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 215
QY 255 CCGCCCTTATCGAATCTGCGAACAACACAGCCTCGATTTGCAATCTACATTTGGCGCAACG 314
Db 216 HisGlyValThrSerAlaProAspThr-----ArgProAlaProGlySerThr 231
QY 315 ACATTTCCGCA---GGCTGCTGTCAACCTCATGTACCTTGGCG-----CAGAAC 362
Db 232 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 251
QY 363 GCGCGCTGCGGTACGCTCTGTTGACGACACACACGCGCGGTGGACCATCTCC 422
Db 252 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 271
QY 423 TGCTCGCCTCG-----ACAGCCATCCCAATATCGAAGTGGCGCTGTCAACCCCT 473
Db 272 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 291
QY 474 TCGTCTTACCAATGGCGCGCATCGCT-----ACC 506
Db 292 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 311
|||
```

```
QY 507 TGACCGACTTCCCGCGCTCAACCGCGGATGCAACAATAATCTTTTACCGCGCAAC 566
Db 312 AlaProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 331
QY 567 GCGGCACCATACTCGCGCGGCAATATCGCGGAGCAATACTTCAAAATCGGTGAGGACA 626
Db 332 AlaProPro-----AlaHisGlyValThrSerAlaProAspThrArgProAla 347
QY 627 CGGTTTTCCGACCTGGACATCTCGCCACCGGACGCGTCTCGCGGAAGTATCGCACG 686
Db 348 ProGlySer-----ThrAlaProProAlaHisGlyValThrSerAlaProAsp 363
QY 687 ACTTCGACCGCTACTGGGCAAGCCATTTCGCGCC-----719
Db 364 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 383
QY 720 ACAACGCCAGCGGCATCTCGCAGCGCAACATCGGCAAGGCTCTTCAAGCACTCGGAT 779
Db 384 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 403
QY 780 ACA-----ACGACGAAACATCCAGAC 800
Db 404 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 423
QY 801 AGCGCTCTGCG-----GCTACCGGAAACCGTCGAAC-----AGTCGCCCTCT 845
Db 424 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 443
QY 846 ACCAAAAAATACAGACGGGACATCGACTGCGCAGGCGTCCAAACCGCGCTGATCAGCG 905
Db 444 ThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVal 458
QY 906 ACACCCCTGCAAAAGGATCGACGCGACCGCGCAACCGCGGATTCGCGGAGGCTGC 965
Db 459 Thr-SerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVa 478
QY 966 AAGACGCGCTCAACAGCGCGGAAAAAGGCTCTATCTGTTTCAACCCCTATTCTGTCCTTA 1025
Db 478 ThrSerAlaProAspThrArgPro-----486
QY 1026 CAAATCCGGCAGACGACCTGGCAAACTGGTCAGGACGGCATAGACGTTTACCGTCC 1085
Db 487 -----AlaProGlySerThrAlaProProAlaHis-GlyValThr-Ser 500
QY 1086 TGACCAACTCGTACAGGCGACG-----ACGTTGCGCGCGTCCATCCCGGCTACGTCA 1139
Db 501 AlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSer 520
QY 1140 ANTACGGAACCGC-----TGCTCAAGCCGGCATCAAACTCTACGAGC 1184
Db 521 AlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSer 540
QY 1185 TGCAACCCCAACCATCGCTCCCGCGCACAAAAGCAAAAGGCGCTGACCGGCA-----1235
Db 541 AlaProAspThrArgProAlaPro-----GlySerThrAlaProProAlaHisGlyVal 558
QY 1236 -----GCTCGTAAACGAGCTGCATCGCAAAACCTTCATG 1271
Db 559 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 578
QY 1272 TGGAGGGAAACGCACTTTCATCGGCTCATTCACCTCGACCCCGTTCGCGACGCG---1328
Db 579 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 598
QY 1329 ---TCAATACCG-----AAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAGAACAGA 1379
Db 599 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 618
QY 1380 TGGAGCGCACTCGCGCGATACACACCGCAATACGCTTACCGGTTACCTTCGCAAAAC 1439
Db 619 ThrSerAlaProAspThrArgProAlaProGlySer---ThrAlaProProAlaHisGly 637
QY 1440 ACAACCGCTGCAATGGCAGCATCCG-----CCACCCGAAAAACCT 1481
```

```
Db 638 ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGly 657
QY 1482 ACCGCAACGAACCGG-----AAGCCAACTTTGGAAACGATCGCCG 1523
Db 658 ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaPro 673

RESULT 31
US-11-096-568A-19864
; Sequence 19864, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19864
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(766)
; OTHER INFORMATION: Ceres Seq. ID no. 12375135
US-11-096-568A-19864
```

```
Alignment Scores:
Pred. No.: 4,49e-08 Length: 766
Score: 233.50 Matches: 148
Percent Similarity: 31.5% Conservative: 49
Best Local Similarity: 23.6% Mismatches: 204
Query Match: 8.2% Indels: 225
DB: 11 Gaps: 31
```

US-10-665-990A-13 (1-1561) x US-11-096-568A-19864 (1-766)

```
QY 160 GGACAACATCTCGCAATTCGGCACACCCCTCATACAACAGCGGTATCCGACATCTACT 219
Db 91 GlyLeuHisPro-----ProArgArgProGlyGlyArgArgLeuArgGlnGlyArg 108
QY 220 GCTCGACGA-----CCCCACGAAGCCCTTGGCCCGCGC 255
Db 109 ValArgArgArgArgArgAlaArgGlyGlyProProArgGlnGlnGlnArg 128
QY 256 CGCCCTTATCAATCTGCCGAACACACGCTCGATTTCGAATACTACTATTGGCGCAACGA 315
Db 129 ValAspProArgLeuHisArgProArgArg-----HisLeu---ArgArg 143
QY 316 CATTTCCGGCAGGTGCTGTTCAACCTCATGTACCTTCGCGCAGAACCGCGGTGCGGT 375
Db 144 HisHisArg-----LeuArgArgHisLeuProProHisGlnGlyArgAlaArg 160
QY 376 ACGCTGCTGTGGA-----CGACAA 396
Db 161 AlaProGlyArgAlaAlaGlnAlaSerGlnArgArgAsnLeuArgHisArgGlnArgArg 180
QY 397 CAACACGCGCGGTGGA-----414
Db 181 AlaHisAlaAlaValGlyGlyHisProArgGlyAspAlaGlyAlaHisGlnHisProPro 200
QY 415 -----CGATCTCTGCTCGCCCTCGACAGCATCCCAATAT 450
Db 201 ProGlyLeuLeuArgHisProLeuArgAspProArgGlyHisHisAlaAlaGlnHis 220
QY 451 CGAAGTCGCGCTGTTCACCCCTTCGTCCTACGCAAAATGGCGCGCACTCGGTACT---507
Db 221 ArgCysGlnProLeu---ProAlaAlaProGlyHisHisArgValGlyArgProGly 239
QY 508 -----GACCGACTTCCCGCGCTCAACCGCGCGCATGCAACAACATCTTTTAC 555
```

```
Db 240 ProAlaLeuLeuHisArgProHisHisGlyProPro----- 252
QY 556 CGCCGCAACCGCCACCATCTCGCGGACGCAATATCTCGCGACGATCTTCAAAGT 615
Db 253 GlnArgAlaGlyArgHisArgArgGlyGlyArgGlyValGlnAsp 272
QY 616 CGGTGAGGACACCGTTTTCG-----CGACCT---GGACATCTCTCGCCACCGG 660
Db 273 Arg---ArgHisArgGlyArgLeuLeuGlnAlaGlnProGlnGlyProArgHisArg 291
QY 661 CAGCCTCGT-----CGCGAAGTATCGCAGCTTCGACCG 696
Db 292 GlnArgHisValArgGlyLeuArgAlaArgGlyHisArgAspValArgGlnArgPro 311
QY 697 -----CTACTCGGCAAGCCATCTCCG----- 717
Db 312 GlyArgProValGlyGlyProValArgArgLeuLeuArgGlyHisGluArgGlnAlaArg 331
QY 718 -----CCACAA----- 723
Db 332 ValHisGlyProProAspProGlnAlaGluAlaProProGlyValHisArgGlyArgGly 351
QY 724 CGCCACGCGCATCTCCG-----CAGCGGCAACATCGCAAGGTCTTCAAGCACTCGG 777
Db 352 HisHisGlyAlaHisProGlyTrpGlnLeuLeuHisGluAlaGlyGlnGluArg 371
QY 778 -----ATACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGCGAAAC 825
Db 372 AlaGlyProAlaAlaGluAlaGlnAlaGlyGlnValArgAlaPro----- 386
QY 826 CGTCGAACAGTCGCGCCCTCTACCAAAAATACAGACGGGCGCATCGACTGCGAGAGCGT 885
Db 387 -----HisValAlaAlaVal 391
QY 886 CCAACCCGCTGATCAGGACACCCC----- 912
Db 392 AlaGlyProProAspArgGlyHisProArgArgHisGlnValHisArgAlaArgGlyGln 411
QY 913 TGCAAAAGACTCGACCG-----CGACCGCGCAAAACCGCGATTGCGGAGGCT 963
Db 412 LeuArgGluArgGlnProGlyHisArgArgProProArg----- 424
QY 964 GCAAGACGCGTCAACAGCCGCGAAAGCGTCTATCTGTGTTTCAACCTATTTCTGCTCC 1023
Db 425 -----GlnGlyAlaAlaArgArgGlnLeuProGlyHisPro----- 436
QY 1024 TACAAATCCGCGACAGCGACTCGGCAAACTGTGTGAGACGGCATAGACTTACCGT 1083
Db 437 -----HisArgValHisGlyGlnArgProProArgHisArgGlnHisArg 452
QY 1084 CTGTACCAACTCGCTACAGGCGACCGACGTTGCGCGCTCCATTCCGCGTACGTCAAATA 1143
Db 453 GlnAlaHisValArgAlaValLeuArgAlaArgGlnArgValLeuGlnArgAlaHis 472
QY 1144 CGGAAACCGTGTCTCAAGCGCGCATCAA----- 1173
Db 473 LeuGlnProGlyArgGlnProGlnProGlnProGlyLeuArgLeuGlnGlyHisArgAsp 492
QY 1174 -----ACTCTAGAGCTGCAACCCACCATCGCTCCCGCCACAAAAGCAA 1221
Db 493 ArgHisGlyLeuLeuLeuArgAlaProValProGlyGlnProHisHisGlnProArg 512
QY 1222 -----AGCCTGACCGGAGCTCCGTAACCAACGCTGCATGCCAAAC 1263
Db 513 AlaGluArgGlyAlaAlaGlnProGlyArgGluLeuProGlyProArgLeuGlyGlnGlu 532
QY 1264 CTTCAATGTGACGGCAA-----ACGCATCTTTCATCGGCTCATT--- 1302
Db 533 AspArgArgGlyAspArgHisProGluAlaHisValValHisLeuHisArgGlyAlaVal 552
QY 1303 -----CAACTCGACCCCGTTCCGACGGCTCAATACCGAAATGGCGTCGTCAT 1353
Db 553 ProGlyArgGlyProAlaProPro---ArgGlyGluHisGlnGlyValGlyGluGluHis 571
```

```
QY 1354 CGAAGAGCCC----- 1362
Db 572 ArgAspProGlyGlnGlnGlyAlaAspHisGluProLeuGlyArgAlaLeuGlnArg 591
QY 1363 -----CAAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCCACCCGAATACGC 1416
Db 592 ProLeuGlnArgGlnGlyAlaAspGlnArgHisArgProArgGlyArgValHisValArg 611
QY 1417 CTACCGGTTACCTTCGACAAACACACCGCTCGAATGGACAGTCCCGCCACCCGAAA 1476
Db 612 GlyGlyArgGlyGlnArgGlnProAlaAlaAspAlaGluAlaAlaArgAlaGlyGly 631
QY 1477 AACCTACCCGAAACCCGAAGC-----CAAACCTTTGGAAACCGATCGC 1521
Db 632 ProArgProGlnGlnArgArgGlyAlaGlyAlaLeuHisValLeuGlnAspHisGln 651
QY 1522 CGCAAAATCTATCCCT 1539
Db 652 ValArgGlyGlyAlaPro 657

RESULT 32
US-11-096-568A-28958
; Sequence 28958, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide.
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28958
; LENGTH: 1480
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1480)
; OTHER INFORMATION: Ceres Seq. ID no. 3095666
US-11-096-568A-28958

Alignment Scores:
Pred. No.: 5,7e-08 Length: 1480
Score: 232.50 Matches: 149
Percent Similarity: 32.5% Conservative: 43
Best Local Similarity: 25.3% Mismatches: 219
Query Match: 8.2% Indels: 180
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-096-568A-28958 (1-1480)
QY 34 CCGCATGAAACACGACGCTCATTTCCCTTTTATGCTCTCTCTCTCTCTCTCTCTCTTC 93
Db 95 ProLysProThrLysProHisProHisProHisProHisProHisProHisProHisPro 112
QY 94 ATGTTGCCCTCCCTGGAAGAACGGACGGAAGCCGTCATTTCATCTTCCAAACCTGT 153
Db 113 ProProHisProLysProArgProHisProLysPro-----ProAsnVal 127
QY 154 CCTCTCGCAACATCTTCGCAATTCGGCA-----CACCCCTCATAAACAACGGGCTATC 207
Db 128 LysProHisProHisProLysProProThrLysProHisProHisProHisProHisPro 147
QY 208 CGACATCTACCTGCTCGACGACCCCAACG----- 236
Db 148 LysHisHisProHisProLysProProThrLysProHisProHisProHisProHisPro 167
QY 237 AAGCCCTTGCG-----CCCGCGCGCCCTTATCGAATCTGCGAACAACAGCCCTCGATT 290
Db 168 LysProProSerThrProLysProProThrThrAsnProPro----- 183
```

```
QY 291 TCGAATACTACATTTGGCGCAGCAGACATTTTCGGCAGGCTGCTGTTCAACCTCATGTACC 350
Db : : : : :
184 ---SerThrProGlnProProThrHisLysProProProCysThrProThr----- 199
QY 351 TTGGCGCAGAACCGCGTGGCGTACGGCTGCTGTGGAGCAGCAACAACACGCGGGT 410
Db : : : : :
200 ---ProProValAlaSerProProMetAla-----ThrProProThrGln----- 213
QY 411 TGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTC AAC 470
Db : : : : :
214 -----MetProProIleAlaThrProProIleAlaLysSerProValAlaThr 229
QY 471 CTTTCGTCTACGCAATGGCGGCACCTCGGTACTGACCGACTCC----- 518
Db : : : : :
230 ProPro-----IleAlaThrProProThrAlaThrProProIle 242
QY 519 -----CCGCTCAACCGCGCATGCACAAATCCT----- 551
Db : : : : :
243 ThrIleProProValAlaThrProProIleThrThrProProIleAlaAsnProProIle 262
QY 552 ---TTACCGCGACAAACCGCCACCATACTCGCGGAGCGCAATATCGGCGAGCAATACT 608
Db : : : : :
263 IleMetProProIleAlaThrProProValAlaAlaProProIleThrAsnProProIle 282
QY 609 TCAAG---TCGTGAGGACACGTTTCGCCGACCTGGACA----- 647
Db : : : : :
283 SerLysProProValThrThrPro-----ProThrThrThrProProIleAlaLysPro 300
QY 648 -----TCCTCGCCACCGGAGCGTCTGCGCGAAGTATCGCAGC 686
Db : : : : :
301 ProIleAlaThrProProIleSerThrProProAlaAlaThrProProAlaAlaThrPro 320
QY 687 ACTTCGACCGTACTGGCAAGCC----- 716
Db : : : : :
321 ProIleThrThrLeuProProAlaLysProProValAlaIleSerProIleValThrPro 340
QY 717 CCCAAGCCGCGGCATATCCGAGCGGAACATCGGCAAGGTCTTCAAGCACTCG 776
Db : : : : :
341 ProValThrProIleAlaGlnProProValAlaThr----- 352
QY 777 GATACAAACGAGAAACATCCAGACGCGCTCTGCGTACCGCGAAGCCGTCGAACAGT 836
Db : : : : :
353 -----ProProThrAlaThrProProValAlaThrProProIleAla 366
QY 837 CGCCCTCTACAAAATAACAGACGGAGCATCGACTGGCAGAGCGTCCAAACCGCGC 896
Db : : : : :
367 ThrProProThrSerLys-----SerProIleSerThrProPro 379
QY 897 TGATCAGCAGACCCCTGTCAAAAGGACTCGACGCGACCGCGCAACCGCGGATGCG 956
Db : : : : :
380 IleSerGluSerProValAlaThrProProThrAlaThrSerProIleLysThrProPro 399
QY 957 GGAGGTGTCAAGAGCGGTCAACACAGCCGCAAAAGCGTCTATCTGGTTTCAACCTATT 1016
Db : : : : :
400 -----ProAlaLysProProValAlaThrProProIle 410
QY 1017 TCGTCCCTACAAATCCGCGCAGACGCTGGCAAACTGTGCGAGGCGGATAGACG 1076
Db : : : : :
411 Ala-----LysSerProIleAlaThr----- 417
QY 1077 TTACCGTCTGACCACTCGCTACAGGCGACGAGTTCGCCGCGTCCATTCGGGTACG 1136
Db : : : : :
418 -----ProProThrAlaThrProProVal---AlaThr 427
QY 1137 TCAAAATACGAAACCGC-----TGCTCAAGCGCGCATCAAACTCTACGAGGTGCAAC 1190
Db : : : : :
428 ProProIleGluLysProProValAlaThrProProThrThrThrProProThrAlaThr 447
QY 1191 CCAACCATCGCTCCCGCGCAAAAGACAAAGGCTGACCGGAGCTCCGTAAACGAGC 1250
Db : : : : :
448 ProProValAlaLysProProValGluThrPro-----ProIleAlaThrProProThr 465
QY 1251 TGCATGCCA-----AAACCTTCATTGTGGAGCGGCAACGCA 1286
```

```
Db 466 AlalysProProIleSerThrProProLysSerProProValAlaThrProProAla 485
QY 1287 TCTTCATCGGTCTATTCAACTCGACCCCGTT-----CCGACGCGC 1328
Db : : : : :
486 AlalysProProIleThrProThrProValLysProProValAlaThrProProLeu 505
QY 1329 TCATATCCGAAATGGCGTGC-----TCATCGAAAGCCCAAAATCG 1370
Db : : : : :
506 AlalysProProValAlaLysProProValValThrProProThrAlaThrProProIle 525
QY 1371 CAGAACAGATGGAGCGACCTCGCGCATACACACCGCAATACCGCTACCGGTTACCC 1430
Db : : : : :
526 AlalysProProIleAlaLysSerProProValAlaThrPro-----ProThrAlaThrPro 543
QY 1431 TCGACA-----AACACACCGCC 1448
Db : : : : :
544 ProValAlaThrProProIleAlaLysProProValValThrProProThrThrPro 563
QY 1449 TGCATGCGCAGATCCCGCCACCGGAAACCTTACCGAAGCCGCAACTTT 1508
Db : : : : :
564 ProThrAlaThrProProValAlaLysProProValAlaThrProProIleAlaThrPro 583
QY 1509 GGAACGATCGCGCAAAATCCTATCCC 1538
Db : : : : :
584 ProThrAlaLysProProIleSerThrPro 593

RESULT 33
US-10-793-626-1000
; Sequence 1000, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1000
; LENGTH: 442
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (442)
; OTHER INFORMATION: variable amino acid
US-10-793-626-1000

Alignment Scores:
Pred. No.: 5,248-08 Length: 442
Score: 232.00 Matches: 89
Percent Similarity: 37.3% Conservative: 57
Best Local Similarity: 22.8% Mismatches: 157
Query Match: 8.1% Indels: 88
DB: 9 Gaps: 9

US-10-665-990A-13 (1-1561) x US-10-793-626-1000 (1-442)
QY 212 ATCTACCTCTCGACAGACCCCGACAGCCCTTGGCCGCGCGCTTATCGAATCT 271
Db : : : : :
129 IleAspLeuPheThrAspGlyHisLysLeuTyGluLysValLeuGluAspIleTyAsn 148
QY 272 GCCGACACAGCCTCGATTTCATCTACATTTGGCGCAACGACATTTCCGCGAGGCTG 331
Db : : : : :
149 AlaGlnAspTyIleHisLeuGluTyThrThrPheGluLeuAspGlyLeuIlyLysArg 168
QY 332 CTGTTCAACCTCATGTACTTGTCCGCGAGAACGCGCGTGTACGCTGTGTGGAC 391
```



```
Db 169 IleLeuAspAlaLeuGluThrLysLeuLysGluGlyLeuGluValLysLeuLysLeuTyrAsp 188
QY 392 GACAAACACACGCGGGTTGGACGATCTCTGCTGCTCGACGACATCCCAATATC 451
Db 189 AspValGlySerLys----- 193
QY 452 GAAGTCGCGCTGTTCAACCCCTTCGCTACGAAATGGCGGCGACTCGGCTACCTGACC 511
Db 194 LysValArgLeuSerLys-----PheLysHisPheArgAlaLeuGlyGlyGluVal 210
QY 512 GAC-----TTCCCC-----CGCTCAACGCGCGCATGCAC 541
Db 211 GluAlaPheProSerLysValProLeuIleAsnPheArgMetAsnAsnArgAsnHis 230
QY 542 AACAAATCTTTACCGCGCAACACGCGCCACCATCTCGGCGGAGCGCAATATCGCGAC 601
Db 231 ArgLysIleIleIleAspGlyGlnIleGlyTyrValGlyGlyPheAsnValGlyAsp 250
QY 602 GAATATCTTCAAGTCGGTGAAGACACCGCTTTTCGCGGACCTGGACATCTCGCCACCGCC 661
Db 251 AspTyrLeuGlyLysLeuGlyTyrTrpArgAspThrHisThrArgValGlnGly 270
QY 662 AGCGTCGCGGCAAGTATCGACAGCTTCGACCGCTACTGGGCAAGCCATTCGCGCCAC 721
Db 271 GluCysIleAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSerGlnSer----- 288
QY 722 AACGCCCGCGCATCTCGCGAGCGCAACATCGCGCAAGGCTCTTCAAGCATCTCGATAC 781
Db 288 ----- 288
QY 782 AACGACGAACATCAGACACGCGCTCTCGCTACTACCGCAAAACCGTCGCAACGTCGCC 841
Db 289 -----HisArgProGlnPheLysPheAspGln 297
QY 842 CTCATACCAAAATACACAGGAGCGATCGACTGGCAGAGCGTCCAAACCGCGCTGATC 901
Db 298 LysTyrPheProLysLysAsnGly-----AspLysGlyAsnAlaIleGlnIleAla 315
QY 902 AGCGACACCTCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGCATTCGCGGGAGG 961
Db 316 SerSerGlyProAlaPheAspLeuHisGln-----IleGluTyrGly 329
QY 962 CTGCAAGACGCGCTCAAAACAGCCGCAAAAGCGTCTATCGTTTCAACCTATTTCGTC 1021
Db 330 TyrThrLysMetIleMetSerAlaLysLysSerIleTyrLeuGlnSerProTyrPheIle 349
QY 1022 CCTACAAATCGGCACAGCAGCTGCGCAAACTGGTTCGAGGACGCGATAGAGCTTACC 1081
Db 350 ProAspGlnSerTyrIleAsnAlaLeuLysMetAlaAlaAsnSerGlyValGluValAsn 369
QY 1082 GTCCTGACCAACTCGCTACAGGCGACCGACGTCGCGCGCTCCATTCGCGGTACGTCAA 1141
Db 370 LeuMetIleProCys-----LysProAspHisProPheValTyrTrpAlaThrPheSer 387
QY 1142 TACCGAAACCGCTGCTCAAAAGCGGCGATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
Db 388 AsnAlaAlaAspLeuAspSerGlyValAsnIleTyrThrTyrGlnAsnGlyPhe---- 406
QY 1202 GTCCCGCGCACAAAGAGCGCTGACCGGCGAGCTCCGTAAACGAGCTGATCGCAAA 1261
Db 407 -----IleHisSerLys 410
QY 1262 ACCTTCATTGTGGACGCAACGATCTTCATCGGCTCATTCACCTCGACCCCGGTTC 1321
Db 411 IleLeuMetIleAspAspGluIleSerSerIleGlySerAlaAsnMetAspPheArgSer 430
QY 1322 GCAGGCTCAATACCGAAATGGGCTGCTCATC 1354
Db 431 PheGluLeuAsnPheGluValAsnAlaPheIle 441
```

RESULT 34

US-11-050-857-487

```
/ Sequence 487, Application US/11050857
/ Publication No. US20060040278A1
/ GENERAL INFORMATION:
/ APPLICANT: Compugen Ltd
/ TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
/ FILE REFERENCE: 1847.1005
/ CURRENT APPLICATION NUMBER: US/11/050.857
/ CURRENT FILING DATE: 2005-01-27
/ NUMBER OF SEQ ID NOS: 1150
/ SEQ ID NO 487
/ LENGTH: 1255
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-050-857-487
```

Alignment Scores:

Pred. No.:	7,01e-08	Length:	1255
Score:	231.00	Matches:	145
Percent Similarity:	35.8%	Conservative:	55
Best Local Similarity:	26.0%	Mismatches:	239
Query Match:	8.1%	Indels:	119
DB:	11	Gaps:	24

US-10-665-990A-13 (1-1561) x US-11-050-857-487 (1-1255)

```
QY 93 CATGTTGCCCCCACTGGAAGACGGAGCGGAAACCGTCATTTCAATACTTTCCAAACCTG 152
Db 436 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 455
QY 153 TCCTCTGGACAATCTCTCAAAATCCGGCACACCCCTCATAAACAAAGGGGTATCCGACA 212
Db 456 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 475
QY 213 TCTACCTGCTGACGACGCCCC---ACGAAGCCC-----TTGCCGCCGCGC 254
Db 476 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 495
QY 255 CCGCCCTTATGATCTGCCGACACACAGCCTCGATTGCAATACATCTATTTGGCGCAACG 314
Db 496 HisGlyValThrSerAlaProAspThr-----ArgProAlaProGlySerThr 511
QY 315 ACATTTCCGGCA---GGCTGCTGTTCAACCTCATGTACCTTGCCG-----CAGAAC 362
Db 512 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 531
QY 363 GCGCGCTGCGGTACGCTGCTGTGGACGACAAACACGCGCGGTGGACGATCTCC 422
Db 532 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 551
QY 423 TGCTCGCCCTCG-----ACAGCCATCCCAATATCGAAGTCGCGCTGTTCAACCCCT 473
Db 552 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 571
QY 474 TCGTCTACCAAAATGGCGGCACTCGCT-----ACC 506
Db 572 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 591
QY 507 TGACCGACTTCCCGGCTCAACGCGCGATGCACAAATCTTTACCGCCGACCAAC 566
Db 592 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 611
QY 567 GCGCCACCATACTCGCGGACGCAATATCGGCGACGAATCTTCAAAGTCGGTGAGACA 626
Db 612 AlaProPro-----AlaHisGlyValThrSerAlaProAspThrArgProAla 627
QY 627 CCGTTTTTCGCGGACCTCGGACATCTCGCCACCGCGAGCGCTCGTGGCGAAGATATCGCACG 686
Db 628 ProGlySer-----ThrAlaProProAlaHisGlyValThrSerAlaProAsp 643
QY 687 ACTTCGACCGCTACTGGGCAAGCCATTTCGCCCC----- 719
Db 644 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 663
```

```
QY 720 ACAACGCCAGCGCATCTCCGACGGCGCAACATCGGCAAGGCTTTCAGCACTCGGAT 779
Db 664 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 683
QY 780 ACA-----ACGACGAAACATCCAGAC 800
Db 684 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 703
QY 801 ACGCGCTCTGCG-----GCTACCGCAAAACCGTCCGAAC-----AGTCGCCCTCT 845
Db 704 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 723
QY 846 ACAAAATAATACAGACGGACGATCGATCGGACAGCGTCCAAACCGCCTGATCAGGG 905
Db 724 ThrArg-----ProAlaProGlySerThrAlaProAlaHisGlyVal 738
QY 906 ACACCCCTCAAAAGACTCGACGGACCGCGCAACCGCGGATTCGGGGAGGCTGC 965
Db 739 Thr-SerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVa 758
QY 966 AAGACGCGTCAAAACAGCCCGCAAAAGGCTCTATCTGTTTTCACCCCTATTTTCGTCCTTA 1025
Db 758 lThrSerAlaProAspThrArgPro----- 766
QY 1026 CAAATCCGCGCAGACGGCACTGGCAAAACTGGTCGACGCGCATAGAGTTACGTC 1085
Db 767 -----AlaProGlySerThrAlaProAlaHis-GlyValThr-Ser 780
QY 1086 TGACCAACTCGTACAGGCGACG-----AGTTGCGCGCTGCATTCGGCTACGTCA 1139
Db 781 AlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSer 800
QY 1140 AATACCGAAACCGC-----TGCTCAAGCCGCGCATCAAACTCTACGAGC 1184
Db 801 AlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSer 820
QY 1185 TGAACCAACCATCGCTCCCGCCCAAAAGACAAAGCCCTGACGGCA----- 1235
Db 821 AlaProAspThrArgProAlaPro-----GlySerThrAlaProAlaHisGlyVal 838
QY 1236 -----GCTCGTAACGACGCTGCATGCCAAACCTTCATG 1271
Db 839 ThrSerAlaProAspThrArgProAlaProAlaProGlySerThrAlaProAlaHisGlyVal 858
QY 1272 TGAACGCAACGATCTTCATCGGCTCATTCACCTCGACCCCGCTCCGACGCG--- 1328
Db 859 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVal 878
QY 1329 ----TCATACCG-----AAATGGGGCTGTCATCGAAAGCCCAAAATCGCAGACAGA 1379
Db 879 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVal 898
QY 1380 TGAAGCGCACCTCGCGGATACACACCGCAATACCGCTACCGCGTTACCTCGCAAAAC 1439
Db 899 ThrSerAlaProAspThrArgProAlaProGlySer---ThrAlaProAlaHisGly 917
QY 1440 ACAACCCCTGCAATGGCAGATCCCG-----CCACCCGAAAAACCT 1481
Db 918 ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGly 937
QY 1482 ACCCGAACGACCG-----ANGCCAACTTTGGAAACGCAATCGCG 1523
Db 938 ValThrSerAlaProAspAsnArgProAlaLeuGlySerThrAlaPro 953

RESULT 35
US-11-043-806-398
; Sequence 398, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; THEREOF FOR DIAGNOSIS OF PROSTATE CANCER
```

```
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 398
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-398
```

```
Alignment Scores:
Pred. No.: 7,01e-08 Length: 1255
Score: 231.00 Matches: 145
Percent Similarity: 35.8% Conservative: 55
Best Local Similarity: 26.0% Mismatches: 239
Query Match: 8.1% Indels: 119
DB: 11 Gaps: 24
US-10-665-990A-13 (1-1561) x US-11-043-806-398 (1-1255)
```

```
QY 93 CATGTTCCCGCCACTGGAAGACGGACGGAAGCGCTCATTTCAATACTTCCAAACCTG 152
Db 436 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 455
QY 153 TCCTCTCGACACATCTGCAATCCGGCACACCCCTCATAAACACGGCTATCCGACA 212
Db 456 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 475
QY 213 TCTACTCTCGACACACCCCGC---ACGAAGGCC-----TTGCCGCCCGCG 254
Db 476 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 495
QY 255 CGCGCTTATCGAATCTCCGAAACACACGCTCGATTCGAATACTACATTTGGCGCAACG 314
Db 496 HisGlyValThrSerAlaProAspThr-----ArgProAlaProGlySerThr 511
QY 315 ACATTTCCGGCA---GGCTGCTGTTCAACCTCATCTACTCTTGCCG-----CAGAAC 362
Db 512 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 531
QY 363 GCGGCGTGGCGTGCCTGCTGTTGGACGACAAACACACGCGGGTTGGACGATCTCC 422
Db 532 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 551
QY 423 TGCTCGCCCTCG-----ACGCCATCCCAATATCGAAGTGCCTGTTCAACCCCT 473
Db 552 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 571
QY 474 TGTCTCTACGCAATGGCGCGCACTCGCT-----ACC 506
Db 572 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 591
QY 507 TGACCGACTTCCCGCCTCAACCGCGCATGCAACAAATTCCTTTACCGCGCAACCC 566
Db 592 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 611
QY 567 GCGCACCATCTCGCGGACCATATCGGCGACGAATCTTCAAAGTCGCTGAGGACA 626
Db 612 AlaProPro-----AlaHisGlyValThrSerAlaProAspThrArgProAla 627
QY 627 CGGTTTTTCGCGACCTGGACATCTCGCACCGCGCGCTCGTCGCGCAAGTATCGCAG 686
Db 628 ProGlySer-----ThrAlaProProAlaHisGlyValThrSerAlaProAsp 643
QY 687 ACTTCGACGCTACTGGGCAAGCCATTCCGCC----- 719
Db 644 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 663
QY 720 ACAACGCCAGCGCATCTCCGACGGCGCAACATCGGCAAGGCTTTCAGCACTCGGAT 779
Db 664 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 683
QY 780 ACA-----ACGACGAAACATCCAGAC 800
```

```
Db 684 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 703
QY 801 ACGCGCTCCTGC-----GCTACCGCGAAACCGTCGAAC-----AGTCGCGCCTCT 845
Db 704 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 723
QY 846 ACCAAAAATACAGACGGAGCGATCGACTGGCAGAGCTCCAAACCCGCTGATCAGCG 905
Db 724 ThrArg-----ProAlaProGlySerThrAlaProAlaHisGlyVal 738
QY 906 ACACCCCTGCAAAAGGACTGACCGCGACCGCGCGCAACCCGCGATTCGCGGAGGTGC 965
Db 739 Thr-SerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVal 758
QY 966 AAGACGCGCTCAAAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1025
Db 758 lThrSerAlaProAspThrArgPro----- 766
QY 1026 CAAATCCGCGACAGACGACCTGGCAAAACTGGTCAGAGCGGATAGACGTTTACCGTCC 1085
Db 767 -----AlaProGlySerThrAlaProAlaHisGlyValThr-Ser 780
QY 1086 TGACCAACTCCTACAGCGGACCG-----ACGTTGCGCGCGCTCCATTCGCGCTAGTCA 1139
Db 781 AlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSer 800
QY 1140 AATACCGAAACCGC-----TGCTCAAAAGCGCGATCAAACTCTACGAGC 1184
Db 801 AlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSer 820
QY 1185 TGCAACCCCAACCATCGCTCCCGCGCACAAAGACAAAGGCGCTGACCGGCA----- 1235
Db 821 AlaProAspThrArgProAlaPro-----GlySerThrAlaProAlaHisGlyVal 838
QY 1236 -----GCTCGTAACAGCGCTGCATCGCATCGCAAACTTCAATG 1271
Db 839 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVal 858
QY 1272 TGGACGGCAACCGCATCTTCATCGCTCATTCATCAACCTCGACCCCGCTTCGCGACGCG- 1328
Db 859 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVal 878
QY 1329 ---TCAATACCG-----AAATGGCGTCTGTCATCGAAAGCCCAAAATCGCAGAACAGA 1379
Db 879 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGlyVal 898
QY 1380 TGGAGCGCACCTCGCGGATACACACCGGATACCGCTACCGGTTACCTCGCAAAAC 1439
Db 899 ThrSerAlaProAspThrArgProAlaProGlySer---ThrAlaProAlaHisGly 917
QY 1440 ACAACCGCTGCAATGGCAGCATCCCG-----CCACCGCGAAACCT 1481
Db 918 ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGly 937
QY 1482 ACCGGAACGACCGG-----AAGCCAAACTTTGGAAACGATCGCGG 1523
Db 938 ValThrSerAlaProAspAsnArgProAlaLeuGlySerThrAlaPro 953
```

RESULT 36

```
US-11-079-463-7814
; Sequence 7814, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
```

```
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7814
; LENGTH: 429
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-7814
```

```
Alignment Scores:
Pred. No.: 6 56e-08 Length: 429
Score: 230.50 Matches: 110
Percent Similarity: 36.9% Conservative: 68
Best Local Similarity: 22.8% Mismatches: 191
Query Match: 8.1% Indels: 115
DB: 11 Gaps: 16
```

US-10-665-990A-13 (1-1561) x US-11-079-463-7814 (1-429)

```
QY 26 ACTATATCCCGCATGAAACACGAGCTCATTTCCCTTTTATGCTCTCTCTCTGTTCA 85
Db 18 SerileGlyThrAlaArgThrGlnAsnileAspSerLeuLeuMet----- 32
QY 86 TGTTCCTTCATGTTGCCCCCACTGGAAAGACGAGCGGAAAGCGCGTCAATTCAATCTTC 145
Db 33 -----GlnArgGluAspThrThrGlnPheIleArgSer 43
QY 146 AAACCTGCTCCTGGACAAACATCTCGCAATCCGGCACACCCCTCATATAACACGGGCTA 205
Db 44 AspSerLeuValLeu---GlnPheLeuGluTyrSerAsnileProIleThrAspAsn--- 61
QY 206 TCGGACATCTACCTGCTCGACGACCCCGCAAGACCCCTTCGCGCGCGCGCGCTTATC 265
Db 62 AsnLysValLysLeuIleLysSerGlyArgGluLysPheGluAspLeuPheGluAlaIle 81
QY 266 GAATCTGCCGAACACAGCCTCGATTTGCAATACTACATTTGGCGCAACGACATTTCCGGC 325
Db 82 ArgGlyAlaLysHisIleHisLeuGluTyrPheAsnPheArgAsnAspSerIleAla 101
QY 326 AGGCTGCTGTTCAACCTCATGTACCTTGGCGGAGACGCGCGCTGCGGTACGCTCGT 385
Db 102 AsnAlaLeuPheAspLeuGlyGluLysValLysGluGlyValLysValAlaGalaMet 121
QY 386 TTGAGACGAC-----AACACACGCGCGGTTGGACGATCTCTGCTCGCGCTC 433
Db 122 PheAspAlaPheGlyAsnTrpSerAsnAsnLysProLeuLysLysArgHisLeuLysAla 141
QY 434 GACAGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCCTACGCAATGGCGC 493
Db 142 IleArgLysGlyIleGluIleValLysPheAspProPhe----- 155
QY 494 GCACCTCGGCTACCTGACGAGCTTCCCGCGCTCAAC-----CGCGCATGCACAAC 544
Db 156 -----LysPheProTyrIleAsnHisAlaAlaHisArgAspHisArg 169
QY 545 AAATCTCTTTACCGCGCACAAACCCGCGCACCATCTCGCGGAGCGCAATATCGCGCAAG 604
Db 170 LysIleAlaValIleAspGlyLysIleGlyTyrThrGlyGlyMetAsnileAlaAspTyr 189
QY 605 TACTTC-----AAAGTCGTGAGGACACCGGTTTTCGCGACCTGGACATCTCTC 652
Db 190 TyrIleAsnGlyLeuProLysIleGly-----ThrTrpArgAspMetHisIleArg 206
QY 653 GCCACCGGAGCGCTCGCGGGAAGTATCGCAGACTTCGACCGCTACTCGGCAAGCCAT 712
Db 207 IleGluGlyAspAlaValAsnileLeuGlnGluIlePheLeuAspIleTrpAsnLysThr 226
QY 713 TCGCGCCCAACAGC-----CACGGCGCATCTCCGCGAGCGCAACATCGCGAAGGTCT 765
Db 227 ThrLysGlnAsnileValGlyGluGluTyrPheProAsnHisProGluArgAlaAspSer 246
QY 766 TCAAGCACTCGGATACAGACGCAAAACATCCAGACACGCGGCTCTCTCGGTACCGGAAAC 825
Db 247 CysAsnThrValIle-----SerIleValAspArgThrPro----- 258
```

```
QY 826 CGTCGAACAGTCGCCCTCTTACCAAAAAAATACAGAGCGGACGCATCGACTGGCAGAGCGT 885
Db 259 -----LysArg 260
QY 886 CCAAACCCGCTGATCAGGCACACCCCTGCAAAAGGACTCGACCGCGCCGCAAAACC 945
Db 261 --AsnSerArgMetLeuSerHisThrTyAlaMet----- 271
QY 946 GCCGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCGCAAAAGGCGTCTATCGT 1005
Db 272 -----SerIleTyAlaAlaGlnHisAspValArgIleVa 283
QY 1006 TTCACCTATTTCGTCCTTACAAAATCCGGCACAGACGACTGGCAAAACCTGGTCAGGA 1065
Db 283 --AsnProTyPheValProThrSerSerIleArgLysAlaLeuLysArgAlaLeuAsnAr 303
QY 1066 CGGCATAGACGTTACCGTCTGACCAACTCGCTACAGGCGCGGACGCTGCGCGCTCCA 1125
Db 303 gGlyThrLysValGluIleMetIleSerSer-----LysSerAspIleProPheThrPr 321
QY 1126 TTCGCGCTACGTCAAATACCGAAACCGCTGCTCAAAGCCGCGCATCAAACTCTACGAGCT 1185
Db 321 oAspAlaSerLeuTyAlaValGlnLysLeuMetLysLysGlyAlaMetIleTyLeuTy 341
QY 1186 GCAACCCCAACATGCGTCCCCCGCCACAAAAGAACGCTGACCGGACGCTCCGTTAAC 1245
Db 341 rAsnGlyGlyPheHis----- 346
QY 1246 CAGCTGTGATGCAAAACCTTCATTGTGGACGGCAACGCATCTTCATCGGCTCATTTCAA 1305
Db 347 -----HisSerLysIleMetMetValAspAspLeuPheCysThrValGlyThrAlaAs 364
QY 1306 CCTGACCCCGCTTCGCGACGCTCAATACCGAAATGGCGCTGCTCATCGCAAAAGCCCAA 1365
Db 364 nLeuAsnSerArgSerLeuArgTyAspTyArgLysThrAsnAlaPheIlePheAspLysAs 384
QY 1366 AATCGCAGAACAGTGGACGCGCACCCCTC---GCCGAT-----ACCAC 1404
Db 384 pileThrGlnGlnLeuAsnAspValPheGluAlaAspMetLeuHisCysThrArgLeuTh 404
QY 1405 ACCCGAA 1411
Db 404 rProGlu 406

RESULT 37
US-11-096-568A-20771
; Sequence 20771, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Theryb
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20771
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(413)
; OTHER INFORMATION: Ceres Seq. ID no. 12387083
US-11-096-568A-20771

Alignment Scores:
Pred. No.: 7.6e-08 Length: 413
Score: 229.50 Matches: 117
Percent Similarity: 40.2% Conservative: 49
Best Local Similarity: 28.3% Mismatches: 153
Query Match: 8.0% Indels: 94
```

```
DB: 11 Gaps: 20
US-10-665-990A-13 (1-1561) x US-11-096-568A-20771 (1-413)
QY 154 CCTCTGGACAAACATCTCTGCAAAATCCGGCACACCCCTCATAAACACGGGCTATCGGACAT 213
Db 18 ProProProArgSerAsnGlnSerGlnSerArgProSerGlnArgArg----- 33
QY 214 CTACTGCTCGACGACCCCGCACGAAGCCCTTGC-----CGCCCGCGCGCCCTTAT 264
Db 34 -----AlaArgGluAsnProHisSerProIleProThrProGlnProPheProPro 51
QY 265 CGAATCTCCCAACACAG-----CCTCGATTTCGAATCTACTACAT 303
Db 52 ArgLeuLeuGlnIleGlnSerSerProSerAlaAsnProProArgLeuGlnIle----- 69
QY 304 TTGGCGCAACGACATTTCCGGCAGGCTGCTGTCAACCTCATGTACCTTGGCGGAGAACG 363
Db 70 -----ProGlyArgGlnAlaAlaAsnGlnProTrpSerAlaSerArgArgSer 85
QY 364 -----CGCGGTGCGGTACGCTGCTGTGGACGACAAACACACCGCGG 408
Db 86 CysSerSerTrpAlaSerSerAlaSerThrAlaSerProArgArgSerSerGluAlaArg 105
QY 409 GTTGCAGCATCT-----CCTGCTCGC-----CCTCGA----- 435
Db 106 ThrAlaArgSerSerArgAlaSerArgProProSerArgCysProSerProArgPheSer 125
QY 436 -----CAGCCATCCCAATATCGAAGTCGCTGTTCACCCCTTCCTCTACGCAA 486
Db 126 SerCysThrSerProSerArgThrArgSerSerProThr-SerThrProThrArgPr 145
QY 487 ATGGCGCGCATCTCGGCTACCTGACCGACTCCCGCGCTCAACCGCGCATGCGACAA 546
Db 145 oTrp-ThrHisPro-----ProArgAlaProThrArgSerAlaAlaCysCys---- 160
QY 547 ATCTTTACCCCGCACACCGCGCCACCATACTCGGCGGACGCAATATCGGCGAGCAATA 606
Db 161 --ProAlaSerProProThrGlySerProCysSerSerSerSerProProThrSerSerA 180
QY 607 CTTCAAGTCGTGAGGACA-----CCGTTTTGCGCGACCTGGACATCCT 651
Db 180 rGSerCysSerArgSerSerProSerProProProCysSerProSerProSerT 200
QY 652 CGCACCAGCGAGCTGTCGGGAAAGTATCGACGACTCG-----ACCGCTACTG 702
Db 200 hrProProSerAlaThrArgSerProSerArgAlaCysSerProSerCysProAlaSerG 220
QY 703 GGCAGGCCATTCCG-----CCCAACGCGCAC 729
Db 220 lYGlyGlySerProProArgSerSerProProSerArgSerSerSerProThrThrProS 240
QY 730 GC-----GCATCATCCGACGCGCAACATCGGCAAGGCTCTTCAAGCACTCGGATACAA 783
Db 240 erProCysSerSerSerSerAlaSerSerProArg----- 252
QY 784 CGAGAAACATCTCAGACACGCGCTCTCGCTACCGCGAAACCGCTGGAACAGT-----CGCC 840
Db 253 --ThrAlaProAlaSerArgAlaCysSer-----ProSerSerProSerP 269
QY 841 CCTCTACCAAAAAATACAGCGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGAT 900
Db 269 roThrCysSerAlaSerSerThrSerAlaSerSerGlyThrSer-Pro----- 284
QY 901 CAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCCCGCATTTGCCGGGAG 960
Db 285 AlaSerSerProCysSerArgThrThrArgAlaSerProProCysAlaArgAlaArgThr 304
QY 961 GCTGCAAGACGCGCTCAACACCGCGAAAGCGCTCTATCTGGTTTACCCCTATTTCGT 1020
Db 305 SerSerGlyGlySerSerProArgValLeuSerSerArgSerThrSerSerSer 324
QY 1021 CCTACAAAATCCGGCACAGCGCACTGGCAAAACTGTGTGAGGACGCGCATAGACGTTAC 1080
```



```
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3385
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3385

Alignment Scores:
Pred. No.: 1,11e-07 Length: 715
Score: 227.50 Matches: 156
Percent Similarity: 33.9% Conservative: 75
Best Local Similarity: 22.9% Mismatches: 229
Query Match: 8.0% Indels: 222
DB: 11 Gaps: 31

US-10-665-990A-13 (1-1561) x US-11-072-512-3385 (1-715)
QY 34 CCCGATGAAACACGACGCTCATTTCCCTTTTATGCTCCTTCTCTGTTTCATGTTCTTC 93
Db ||| : : : : : ||| |||
QY 4 ProLysArgSerSerLysProSerMetSerLeuAlaProSerGlySerSerMet----- 21
Db ||| : : : : : ||| |||
QY 94 ATGGTTGCCGCCACCTGGAAGAACGGACGGAAGCCG-----TCATTTCATATCTCCAA 147
Db ||| : : : : : ||| |||
QY 22 -----ProThr-----AlaAspProLysProProAlaSerLeuLysSerThrLys 36
Db ||| : : : : : ||| |||
QY 148 ACCTGCTCCTCTGGACACAT-----CCTGCAAAATCCGGCA----- 183
Db ||| : : : : : ||| |||
QY 37 SerAlaThrProAsnArgSerLeuValProThrLysProAlaThrSerArgAsnSerVal 56
Db ||| : : : : : ||| |||
QY 184 ---CACCCCTCATAAACACGGGCTATCCGACATCTACCTGCTCGACGCCGCCACGAAAGC 240
Db ||| : : : : : ||| |||
QY 57 MetSerProSerSerLysSerThrLysSerThrSerThrLysArgAlaProSerAsn 76
Db ||| : : : : : ||| |||
QY 241 CTTGCCGCCGCCGCCCTTTATCGAATC-----TGCCGAACACAGCCTCGATTT 291
Db ||| : : : : : ||| |||
QY 77 -----ArgProSerSerArgSerArgValArgSerLysAlaArgThrProSerArgVal 94
Db ||| : : : : : ||| |||
QY 292 GCAATACATATTCGGCGAACGACATTTCCGGGAGGCTGCTGTTCAACCTCATGTACT 351
Db ||| : : : : : ||| |||
QY 95 Ser-----ThrAspThrArgThrSerLysAlaSerLysAlaSerAspValArg 110
Db ||| : : : : : ||| |||
QY 352 TCCCGCAGAACCGG-----CGTGGCGGTAGCCTGCTGTTGGACGACAAACAA 399
Db ||| : : : : : ||| |||
QY 111 CysHisGlnArgArgGlyThrHisSerArgGlyArgThrPro-----GlyArgArgGly 128
Db ||| : : : : : ||| |||
QY 400 CACGCGCGGTTGGACGATCTCCTCGCTCGACAGCCATCCCAATATCGAAGTCG 459
Db ||| : : : : : ||| |||
QY 129 SerArgSerLysArgSerProSerArgAlaSerThrProGlyArgIleArgThr--- 147
Db ||| : : : : : ||| |||
```

460 CCTGTTCAACCCCTTCGTCCTACGCAAAATGGCGGCACCTCGGCTACCT----- 507
148 -----HisGlyAlaArgProGlyMetAlaSerArgValArgThrProThrSerGlnGln 165
507 ----- 507
166 LysGlySerArgGlyLysSerTyrGlyArgProArgThrSerAsnArgGluArgSerAsp 185
508 -----GACCGACTTCCCGGCC----- 524
186 SerGlnProArgAsnLeuSerLysLysSerTyrArgProGlyGlySerGlyIleGly 205
525 -----TCAACCCCGCATGCACACAAATCCCTTACCG 557
206 ArgSerSerGluLeuAlaValThrProSerThrAlaLysCysGlnThr----- 221
558 CCGACAAACCGGCGCCACCATCTACGCGGACCAATATCGCGGACGAATACTTCAAAGTCG 617
222 ProThrGlyIleProSerLysGluLysSerAspAsnProSerProSerSerArgLys 241
618 GTGAGGACA-----CCGTTTTCGCCACCTGG 644
242 VallysSerTyrGlyGlnMetIleProSerArgGluLysSerTyrSerProThrGlu 261
645 ACATCCTCG-----CCACCGGACGCTCGTCGGCGAAGTATCGACGACTTCCACCGCTACT 701
262 MetSerSerArgValLysSerTyrAsnGlnAlaSerThrArgSerArgProGlnSerHis 281
702 GGGCAAGCCATTCCGCGCACCAACCGCCATCATCCGCGGCGGCAACATCGGCAAGG 761
282 SerGlnSerArgSerProArgSerArgSerGlySerGlnLysArgThrHisSerArg 301
762 GTCTTCAAGCACTCG----- 776
302 ValArgSerHisSerTrpLysArgAsnHisSerArgAlaArgSerArgThrArgLysGly 321
777 -----GATACAACGACGAAACATCCACACGCGCTCCTCGCTACCGCGAAACCGTCG 830
322 IleLeuSerGlnMetGlyArgHisSerGlnSerArgSerHisSerLysGlySerGln 341
831 AAC----- 833
342 AsnGlnSerArgThrProArgArgGlyArgSerHisAsnTrpSerArgAsnProSerLys 361
834 -----AGTCGCCCTCTACCAAAAAAATACACGCGGACGCTCGACT 875
362 GluArgSerHisSerHisSerArgSerSerLysGlu-----ArgAsp-HisArgGly 379
876 GGCAGAGCGTCCA-----AACCGCGCTGATCAG-- 903
379 yserSerSerProArgLysGluSerGlyArgSerGlnSerGlySerProAsnLysGlnAr 399
904 -----CGACACCCCTGCAAAAGGACTCGACCGGACCGCGCAACCGC 947
399 gAspHisSerArgSerArg-SerProAsnLysAlaArgAspArgSerArgSerArgSerP 419
948 CGATTGCCGGAGGCTGCAAGCGCTCAACAGCCGCGGAAAAAGCGTCTATCTGGTTT 1007
419 roTyrLysAlaArg-AspArgSerArgSerArgSerProAsnLysAlaArgAspCysSer 438
1008 CACCTATTTCGTCCTACAAAATCCGCGCACAGCGACTGGCAAACT-----GGTGC 1061
439 ArgSerArgSer-ProTyrLysAlaArgAspArgSerArgSerArgSerArgSerArgSe 478
1062 AGGACGGCATAGAGTTTACCGTCTCTGACCAACTCGCT----- 1098
458 aArgAspHisSerArgSerArgSerProAsnLysAlaArgAspArgSerArgSerArgSe 478
1099 -----ACAGCGGACCG 1109
478 rProSerLysGluArgAspHisSerGlnLeuGlySerProSerLysGluArgAspHisAr 498
1110 ACGTTGCCGCGCTCCATTTCGCGCTACGTCAAAATACCGGAAAAACCGCTGCTCAAGCGCGCA 1169

```
Db 498 gArgSerArgSerProSerLysGluArgGlnCysArgGlnSerArgSerSerLysG1 518
QY 1170 TCAAACTCTACGAGCTGCAACCCACCATCGCGTCCCGCCACAAAACACAA- 1221
Db 518 uArgAspHisArgArgSerProSerLysGluArgGlnArgGlnSerArgSe 538
QY 1222 -----AGCCTGACCGGAGCTCCGTAACACGAGCTGCATGCCAAAACCTTCATTG 1271
Db 538 rProAsnLysGluArgAspArgSerGlnSerArgSerProSer 552
QY 1272 TGGACGGCAACCGCATCTTCATCGGCTCATCAACCTCGACCCCGTTCCGACGGCTCA 1331
Db 553 -----GluGluArgGluHisArg-----GlnSerArgSerPro-----SerLy 565
QY 1332 ATACCGAAATGGCGTCTCATCGAAAGCCC-----CAAAATCGCAGAACAGA 1379
Db 565 sGluArgAspArgArgGtrpArgSerProSerLysGluArgGluArgGlnSerAr 585
QY 1380 TGGACGCAACCTCCCGCATACACACCCGAATACGCTCAGCGCT-----TA 1427
Db 585 gSerSerGluGluArgAspHisSerArgSerArgSerProAsnLysGlnSerGlyTy 605
QY 1428 CCCTCGA-----CAACACAAACCGCTGCAATGCGACGATCCCGCCACCCGAAAAA 1478
Db 605 rSerArgProArgAlaSerSerLysGluLysAlaHisSerArgSerArgThrProSerLy 625
QY 1479 CCTACCCCAACCAACCCGCAACCTTTGGAACGATCGCGCAAAATCCT 1533
Db 625 s-----GluGlyAsnHisSerGlnSerArgThrSerSerLysGluSerAspPro 641
```

RESULT 40

```
US-10-510-386-238
; Sequence 238, Application US/10510386
; Publication No. US20050244922A1
```

GENERAL INFORMATION:

```
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Joergensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 238
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-238
```

Alignment Scores:

```
Score: 1.2e-07 Length: 396
Pred. No.: 226.50 Matches: 97
Percent Similarity: 35.8% Conservative: 58
Best Local Similarity: 22.4% Mismatches: 177
Query Match: 7.9% Indels: 101
DB: 9 Gaps: 11
```

US-10-665-990A-13 (1-1561) x US-10-510-386-238 (1-396)

```
QY 206 TCCGACATCTACCTGCTCGACGACCCCAAGCCCTTGCCTCGCGCGCCCTTATC 265
Db 40 SerAspIleGluLeuIleHisAsnGlyGluAspLeuCysGluArgLeuLeuAspIle 59
QY 266 GAATCTGCCGAACACAGCTCGATTGCAATPACTACATTTGGCGCAACGACATTTCCGCG 325
Db 60 ArgGlnAlaGluSerSerValHisValMetPheThrIleValLysAsnAspIleSer 79
QY 326 AGGCTGCTGTTCAACCTCATGCTTGCCTGCGCAGAACCGCGCTGCGCTACGCTGCTG 385
```

```
Db 80 LeuGluPheLeuLysValLeuLysAspLysAlaLysSerGlyValCysValArgLeuLeu 99
QY 386 TTGACGACACAACAACACGCGCGGTGGACGATCTCTGCTCGCCCTCGACGACCATCCC 445
Db 100 IleAspArgIleGlyAlaMetLysValLysLysThrLeuSerGlyLeuLysGlnSer 119
QY 446 AATATCGAAGTGCCTGTTCAACCCCTTCCTCTAGCAATGGCGGCACTCGGCTAC 505
Db 120 GlyValHisValPhePheAlaAsn-----LysProGlyPhe 131
QY 506 CTGACCGACTTCCCGCTCAACCGCGCATGACACAACAAATCTTTACCGCGCAACAC 565
Db 132 ProTyrPhePheTyrArgLeuAsnAlaArgAsnHisArgLysIleAlaValIleAspGly 151
QY 566 CGCGCCACCATCTCGCGGAGCAATATCGCGACCAATACTTC-----AAAGTGGGTGAG 622
Db 152 LysIleGlyTyrValGlyGlyPheAsnIleAlaLysGluTyrLeuGlyLysAlaGlu 171
QY 623 GACACCGTTTTCGCGGACCTGGACATCTCGCCACCGCAGCGTCTCGGCGAAGTATCG 682
Db 172 PheGlyProTyrLysAspTyrHisLeuArgMetThrGlyGluGlyValAlaAspLeuGln 191
QY 683 CAC-----GACTTCGACCGCTACTTGGCAAGCCATTCGCGCCACCAACGCCACG 730
Db 192 HisIlePheIleSerAspPheLys----- 199
QY 731 CGCATCATCGCAGCGCAACATCGGCAAGGCTTTCAGCACTCGGATACACGACGAA 790
Db 199 ----- 199
QY 791 ACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCCGAAACAGTCCGCC- 841
Db 200 -----ArgGluAlaProGlnAlaLysProAlaAsnSer 210
QY 842 CTCTACCAAAAATACAGCGGAGCATCGACTGGCAGAGCGTCCAAAACCGCTGATC 901
Db 211 ValPheProLeuGlnGlnGlyAlaValThrHisThrThrHisAlaThrLysGlyPhe 230
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGCGCAGCCGCAACCGCGATTCGCGGAGG 961
Db 231 Ser-----LeuGluGluLys 235
QY 962 CTGCAAGACGCGCTCAAAACAGCCGAAAAAGGCTCTATCTGTTTCCACCTATTTCGT 1021
Db 236 TyrIleSerPheIleGluGlnAlaLysGluArgIleMetIleCysThrProTyrTyrIle 255
QY 1022 CCTACAAAATCCGCGCACAGACGCTCGGCAAAACTGGTGGAGGACGATAGAGTTACC 1081
Db 256 ProSerProAlaLeuGlnGlnAlaValLeuSerAlaArgGluArgGlyValIleValSer 275
QY 1082 GTCTGTACCAACTCGCTACAGGCGACCGAGCTTCCCGCCCTCCATTCCGCGCTAGTCAA 1141
Db 276 ValLeuValPro-----MetLysProAspHisProLeuValLysGluAlaAlaTyrThr 293
QY 1142 TACCGAAAAACCGTGTCTCAAAAGCGCGCATCAAACTCTACAGAGTGTCAACCAACCATGCC 1201
Db 294 HisPheProAlaLeuLeuLysAlaGlyCysTyrIleTyrArgTyrTyr----- 309
QY 1202 GTCCCGCCCAAAAAGCAAAAGCGCTGACCGGCGAGCTCCGTAACCAACCGCTGATGCCAAA 1261
Db 310 -----ArgGlyPhe-----TyrHisAlaLys 316
QY 1262 ACCTTCATCTGAGCGCAACGCATCTTCTCGGCTCATTCACCTCGACCCCGCTTCC 1321
Db 317 AlaLeuIleValAspAspArgHisValMetIleGlyThrSerAsnPheAspAsnArgSer 336
QY 1322 GCACGGTCTAATACCGAAATGGCGCTCGTCTCATCGAAAAGCCCAAA----- 1366
Db 337 LeuPheLeuAsnAspGluValAsnValIleHisAspLysAspTyrThrLysGlnPhe 356
QY 1367 -----ATCGCAGAACAGATGGAGCGCAGCCCTCGCCGATACCACACCGCAATAGCC 1417
Db 357 PheAspValValLysGluSerIleGluHisAlaGluLeuLeuThrLysGluArgTyrAla 376
```

QY 1418 TACCGGTTACCTCGACAAACACACACCGCCTGCAATGG 1456
Db 377 LysArgProValMetGlnArg-----ProValGluTrp 387

Search completed: May 2, 2006, 06:05:44
Job time : 124.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:02:04 ; Search time 76 Seconds
(without alignments)
2898.236 Million cell updates/sec

Title: US-10-665-990A-13
Perfect score: 2852
Sequence: 1 caaatacaggaatgcgt.....tgccatcgagggtttatta 1561

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10665990/runat_01052006_111941_9494/app.query.fasta_1
-DB=UniProt -QWTF=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10665990 @CGN 1.1 466 @runat_01052006_111941_9494 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DRV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -DELOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot.sprot.*

2: uniprot.trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2689	94.3	525	Q6W764_NEIGO	Q6W764 neisseria 9
2	2681	94.0	525	Q5F887_NEIG1	Q5F887 neisseria 9
3	2630	92.2	525	Q9JTT1_NEIMA	Q9JTT1 neisseria m
4	2573	90.2	508	Q9JYU0_NEIMB	Q9JYU0 neisseria m
5	1182	41.4	524	Q7VUJ6_BORPE	Q7VUJ6 bordetella
6	1181	41.4	492	Q7WP63_BORBR	Q7WP63 bordetella
7	1104.5	38.7	493	Q8FIS6_ECOL6	Q8FIS6 escherichia
8	1102.5	38.7	493	1YND_CECOL1	P75919 escherichia
9	1102.5	38.7	493	Q3RUS5_SHIFL	Q3RUS5 shigella fl
10	1100.5	38.6	493	Q7AFAB_ECO57	Q7AFAB escherichia
11	1098	38.5	528	Q57QL0_SALCH	Q57QL0 salmonella
12	1093.5	38.3	493	Q8X917_ECO57	Q8X917 escherichia
13	1092.5	38.3	514	Q9CKM5_PASMU	Q9CKM5 pasteurella
14	1089.5	38.2	494	Q8ZQ28_SALTU	Q8ZQ28 salmonella
15	1089.5	38.2	495	Q8Z7M0_SALTI	Q8Z7M0 salmonella
16	1089.5	38.2	495	Q5PGY9_SALPA	Q5PGY9 salmonella

17	940.5	33.0	519	2	Q6F8A6_AC1AD	Q6F8A6 acinetobact
18	881.5	30.9	516	2	Q4LR64_9BURK	Q4LR64 burkholderi
19	881	30.9	505	2	Q87JB2_VIBPA	Q87JB2 vibrio para
20	872.5	30.6	501	2	Q8D641_VIBVU	Q8D641 vibrio vuln
21	869.5	30.5	524	2	Q8D861_VIBVU	Q8D861 vibrio vuln
22	863.5	30.3	554	2	Q7MMB3_VIBVY	Q7MMB3 vibrio vuln
23	862.5	30.2	501	2	Q8D800_VIBVU	Q8D800 vibrio vuln
24	852	29.9	570	2	Q4FTL4_9GAMM	Q4FTL4 psychrobact
25	850.5	29.8	570	2	Q8PIM4_XANAC	Q8PIM4 xanthomonas
26	849.5	29.8	520	2	Q6MQL7_BDEBA	Q6MQL7 bdellovibri
27	846	29.7	464	2	Q5QUY8_IDILO	Q5QUY8 idiomarina
28	843	29.6	520	2	Q4UWU7_XANCP	Q4UWU7 xanthomonas
29	843	29.6	520	2	Q8P7A5_XANCP	Q8P7A5 xanthomonas
30	828	29.0	523	2	Q6G185_BAROU	Q6G185 bartonella
31	825	28.9	523	2	Q6G4X0_BARHE	Q6G4X0 bartonella
32	812.5	28.5	466	2	Q98NZ3_RHILO	Q98NZ3 rhizobium 1
33	807	28.3	518	2	Q8UEX3_AGRIT5	Q8UEX3 agrobacteri
34	804.5	28.2	466	2	Q984B3_RHILO	Q984B3 rhizobium 1
35	794	27.8	520	2	Q6F853_AC1AD	Q6F853 acinetobact
36	792	27.8	529	2	Q9HTP4_PSEAE	Q9HTP4 pseudomonas
37	774	27.1	550	2	Q62E43_BURWA	Q62E43 burkholderi
38	773	27.1	517	2	Q88CA5_PSEPK	Q88CA5 pseudomonas
39	772.5	27.1	540	2	Q4LJD9_9BURK	Q4LJD9 burkholderi
40	769	27.0	542	2	Q88BC2_PSESM	Q88BC2 pseudomonas
41	766	26.9	550	2	Q63J33_BURPS	Q63J33 burkholderi
42	762.5	26.7	510	2	Q6LII3_PHOPR	Q6LII3 photobacter
43	755.5	26.5	526	2	Q4ZZX0_PSESY	Q4ZZX0 pseudomonas
44	751	26.3	515	2	Q5E616_VIBF1	Q5E616 vibrio fisc
45	750	26.3	521	2	Q92QQ3_RHIME	Q92QQ3 rhizobium m

ALIGNMENTS

RESULT 1
Q6W764_NEIGO
ID Q6W764_NEIGO PRELIMINARY; PRT; 525 AA.
AC Q6W764;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DT Phospholipase D (Fragment).
GN Name:pld;
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1291;
RX MEDLINE=22935225; PubMed=14573659;
RX DOI=10.1128/IAI.71.11.6381-6391.2003;
RA Edwards J.L., Entz D.D., Apicella M.A.;
RT "Gonococcal phospholipase d modulates the expression and function of
RT complement receptor 3 in primary cervical epithelial cells.";
RL Infect. Immun. 71:6381-6391(2003).
DR EMBL; AY307929; AAQ77232.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
FT NON TER 525
SQ SEQUENCE 525 AA; 55224 MW; 4B626F5264E076D1 CRC64;

Alignment Scores:
Pred. No.: 6.33e-142 Length: 525
Score: 2689.00 Matches: 520
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 94.3% Indels: 0
DB: 2 Gaps: 0

Db 6 LysileGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer 25
QY 62 CTTTATGCTCTCTCTGTTCTCATGTTCTTCCATGGTGGCCCTCATCGAGCGGACG 121
Db 26 LeuLeuCyLeuLeuLeuCySerCySerSerIleProLeuProLeuGluGluArgThr 45
QY 122 GAAGCGCTCAATTCATCTTCCAAACCTGCTCTCTCGAGCAACATCTGCAATCCGG 181
Db 46 GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnLeuGlnIleArg 65
QY 182 CACACCCCTCATACACAGCGCTATCCGACATCTACCTGCTCGAGGACCCCGAGGCC 241
Db 66 HisthrProHisAsnAsnGlyLeuSerAspIleTyrLeuLeuAspProHisGluAla 85
QY 242 CTGCGCGCGCGCGCGCTTATCGAATCTGCGCAACACAGCTCGATTTCGAATCTAC 301
Db 86 PheAlaAlaArgAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCGAACAGCATTTCCGCGAGGCTGCTGTTCACCTCATGTACTTTCGCGCAGAA 361
Db 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaGlu 125
QY 362 CGCGCGTCCGCTACGCTGCTGTGGACGACAAACACAGCGCGGGTTGGAGCATCTC 421
Db 126 ArgGlyValArgValArgLeuLeuAspAspAsnAsnThrArgGlyLeuAspLeu 145
QY 422 CTGCTCGCGCTCGACAGCATCCCAATATCAAGTGGCGCTTCAACCCCTTCGCTCTA 481
Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATGCGCGCACTCGGCTACTCGACCGACTTCCCGCGCTCAACCGCGCATGCAC 541
Db 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMetHis 185
QY 542 AACAAATCTTTACCGCGCAACACCGCGCACCATCTCGCGGACGCAATATCGGCGAC 601
Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAAGTCGGTGAGCACCGTTTTTCGCGCACTCGGACATCTCGCCACCGGC 661
Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTCTGCGCGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCATTTCGCGCCAC 721
Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACCGCAGGCATCATCCGCGGCAACATCGCGAAGGCTCTTCAAGCACTCGGATAC 781
Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGACGAAACATCCAGACGCGCTCTCGGCTACCGCGAAGCGTTCGACAGTCGCGCC 841
Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCTACCAAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAAACCGGCTGATC 901
Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGGACCGCGCAACCGCGATTCGCGGAGG 961
Db 306 SerAspSerProAlaLysGlyLeuAspArgAspArgLysProProlleAlaGlyArg 325
QY 962 CTGCAAGCGCTCAACACGCCCGCAAAAGCGTCTATCTGTTTCACCTATTTCGTC 1021
Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CCTACAAAATCCGCGCACAGCGCACTCGCAAAATCGTGTGAGGACGATAGAGTTTACC 1081
Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCTGACCAACTCGCTACAGGCGACCGAGTTCGCGCGCTTCCATTCGCGCTACGTCAAA 1141

Db 366 ValLeuThrAsnSerSerLeuGlnAlaThrAspValAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAAACCGCTCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACATGCC 1201
Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
QY 1202 GTCCCCCGCCAAAAGACAAAGGCTGACCGGACGCTCCGTAACACGAGCTCGATGCCAAA 1261
Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTGAGCGCAACGCGATCTTCATCGGCTCATTCACCTCGACCGCGCTTCC 1321
Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCAGGCTCAATACCGAAATGGGCTGTCATCGAAGCCGCCAAAATCGCAGACACATG 1381
Db 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGCATACACACCCGAATACCGCTACCGGTTACCTCGCAACACAC 1441
Db 466 GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis 485
QY 1442 AACCGCTCGAATGGCAGCATCCCGCACCGCAAAAACCTACCGCAACCGCAAGGCC 1501
Db 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAATTTGGAAACGCGATCGCGCAAAAATCTATCTCTGCTGCGCATCGAAGTTTATTA 1561
Db 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu 525

RESULT 3
Q9JTT1_NEIMA
ID Q9JTT1_NEIMA PRELIMINARY; PRT; 525 AA.
AC Q9JTT1,
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=NMA1646;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Pratt B.G., Harrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR ENBL; AL162756; CAB84874.1; -; Genomic_DNA.
DR PIR; B81859; B81859.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 525 AA; 1CFB7AC5B82F1B02 CRC64;

Alignment Scores:
Pred. No.: 1,24e-138 Length: 525
Score: 2630.00 Matches: 509
Percent Similarity: 98.7% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 7
Query Match: 92.2% Indels: 0

```
DB: 2 Gaps: 0
US-10-665-990A-13 (1-1561) x Q9UTTI_NEIMA (1-525)
QY 2 AAAATACAGGCAATCCGCTCTGAAACTATATCCCGATGAAAAACACGACGCTCATTTCC 61
DB 6 LysThrGlnAlaMetProSerGluThrIleSerLeuMetLysThrArgSerLeuIleSer 25
QY 62 CTTTATGCTCCTCTCTCTGTTCAATGTTCTTCATGTTGCCCTCCCTACCTGAAAGACGGAGC 121
DB 26 LeuLeuCysLeuLeuLeuCysSerCysSerTrpLeuProProLeuGluGluArgThr 45
QY 122 GAAAGCCCTCATTTCAATACTTCCAAACCTGCTCCTCTGGACACATCTGCAATCCGG 181
DB 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTCATAAACAGGGGTATCCGACATCTACTCTGCTCGACGCCCCACCAAGGCC 241
DB 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
QY 242 CTTGCGCGCGCGCTTATCGAATCTGCGGAACACAGCTCGATTTGCAATPACTAC 301
DB 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGTACCTTCCCGCAGAA 361
DB 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
QY 362 CGCGCGGTGCGGTACGCTGCTGTGTGGACACAAACACGCGGGGTGGACGATCTC 421
DB 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCAACCCCTCTGCTCTA 481
DB 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATGCGCGGCTACCTGCTACCGACCTTCCCGCGCTCAACGCGCGCATGCAC 541
DB 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMetHis 185
QY 542 AACAAATCTTTACCGCGCACAAACCGCGCACCATCTCGGGGACGCAATATCGGCAC 601
DB 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAAGTCGGTAGACACCGTCTTCGCGCCTGACCTGGAACATCTCGCACCGGC 661
DB 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTCGTCGGGAAGTATCGACGCTTCCGACCGCTACTGGGCAAGCATTCGCGCCAC 721
DB 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCAGCGCATCATCCGACGCGCAACATCGGCAAGGGTCTTCAAGCACTCGGATAC 781
DB 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGAGAAACATCCAGACACCGCTCTCGCGCTACCGCGAAACCGTCGAAACAGTCGCGC 841
DB 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTTCTACCAAAAATACAGCGGCGCATCGACTGCGAGCGGTCCAAACCGCTGATC 901
DB 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGGACTCGACGCGGACCGCGCAACACCCCGCATTCGCGGGAGG 961
DB 306 SerAspAspProAlaLysGlyLeuAspArgAspArgArgLysProIleIleArg 325
QY 962 CTGCAAGACGCGCTCAAAACAGCCCGGAAAGCGTCTATCTGGTTTACCCCTATTTCTGTC 1021
DB 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CCTCAAAAATCCGGCACAGCCACTTGGCAAAACTGGTGCAGGACGGCATAGACGTTACC 1081
```

```
DB 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCTGACCAACTCGCTACAGCGCAGCGAGCTTCCCGCGCTCCATTCCGGCTACGTCAAA 1141
DB 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAAAACCGTGTCTCAAAAGCCGGCATCAAACCTTACGAGCTGCAACCAACCATGCC 1201
DB 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
QY 1202 GTCCCGCGCCACAAAGACAAAGGCTGACCGGAGCTCCGTAACACAGCTCATGCCAAA 1261
DB 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTGGACGGCAAAACGCGATCTTCATCGGCTCATTAACCTCGACCCCGTTCC 1321
DB 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCAGCGCTCAATACCGAAATGGCGTGTCTATCGAAAGCCCCAAAATCGGAGAACAGATG 1381
DB 446 AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGCATACCCACACCGCAATACCGCTACCGCTTACCTCGACAAACAC 1441
DB 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
QY 1442 AACCGCTCGCAATGGCACGATCCCGCCACCCGAAAAACCTACCCGAACGAACCCGAAGCC 1501
DB 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTGGAAACGCGATCGCGGAAAAATCCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
DB 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525

RESULT 4
Q9JYU0_NEIMB PRELIMINARY; PRT; 508 AA.
ID Q9JYU0_NEIMB PRELIMINARY;
AC Q9JYU0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthetase family protein.
GN OrderedAccession=NM01434;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA Tetelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Citron H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Masignani V., Pizza M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002098; AAF41795.1; -; Genomic_DNA.
DR PIR; B81083; B81083.
DR TIGR; NMB1434; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
```



```
QY 1490 GAACCCGAAGCCAAACTTTGGAAACGCATCGCCGCAAAATACTTATCTCTGCTGCGCCATC 1549
Db 470 GluProAlaThrSerPheThrIpsArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGCTTTATTA 1561
Db 490 GluTrpLeuLeu 493

RESULT 9
Q83RU5 SHIFL
ID Q83RU5_SHIFL PRELIMINARY; PRT; 493 AA.
AC Q83RU57_QYUCY8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Putative synthase.
GN Name=yndC; OrderedLocusNames=S1116, SF1042;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE005674; RAN42664.2; -; Genomic DNA.
DR EMBL; AE016981; AAP16548.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 493 AA; 55912 MW; 495604AD934C6B92 CRC64;
```

Alignment Scores:

Pred. No.:	2,32e-53	Length:	493
Score:	1102.50	Matches:	221
Percent Similarity:	65.5%	Conservative:	83
Best Local Similarity:	47.6%	Mismatches:	143
Query Match:	38.7%	Indels:	17
DB:	2	Gaps:	5

US-10-665-990A-13 (1-1561) x Q83RU5_SHIFL (1-493)

```
QY 200 GGGCTATCCGATCTACTCTGCTGACGACCCCGCCAGAGCCCTTGGCGCGCGCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluIysSerLeuAlaPheAlaAlaArgTyrArg 56
QY 260 CTTATCGAATCTGCCGAACACAGCTCGATTGCAATACATATTGGCGCAACACATTT 319
Db 57 LeuAlaGluMetAlaGluHisThrLeuAspValGlnTyrIleTrpGlnAspMet 76
```

```
QY 320 TCCGCGAGGCTGCTGTTCACACCTCATGTACCTTCCGCGAGAACCGGCGTGGCGTAGCG 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTGGACGACAAACACGCGGGTGGACGATCTCCTGCTCGCCCTCGACAGC 439
Db 97 LeuLeuLeuAspAsnAsnThrProGlyLeuAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCGAAGTGGCGCTGTTCACACCCCTTCCTCCGCAAAATGGCGCAGCTC 499
Db 117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGCTACCTGACCGATCTCCCGCTCAACCGCGCATGACACAAATCTTACCGCC 559
Db 137 GlyTyrIleThrAspPheSerArgLeuAsnArgMetHisAsnLysSerPheThrVal 156
QY 560 GACAAACCGCGCACCATCTCGCGGAGCAATATCGCGACGAATATCTTCAAAAGTCGGT 619
Db 157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
QY 620 GAGGACACCGTTTTCGCGACATCTCGCCACCGCGCAGCGCTCGTCGGCGAAGTA 679
Db 177 GluGluProLeuPheSerAspLeuValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCACGATCTGACCGCTACTGGGCAAGCCATTCGCGCCACACCGCCGCGCATATC 739
Db 197 AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnGlnValLeu 216
QY 740 CGCAGCGCAACATCGCGAAGGT-----CTTCAAGCACTCGGA 778
Db 217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGACGAAACATCCAGACACGCGCTCTCGCGCTACCGCGAAACCGTCAACAGTCG 838
Db 234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAAATAACAGACGGGACGATCGATCGGAGAGCGTCCAAACCGCGCTG 898
Db 250 PropheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCAGACCCCTGCAAAAAGGACTCGACCGCAGCCGCGCAACCCCGCATTCGCGG 958
Db 270 LeuSerAspAspProAlaLysGlyGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGACGCGCTCAAAACAGCCGCAAAAGCGTCTATCTGGTTTCACTATTTC 1018
Db 290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleIleSerSerTyrPhe 309
QY 1019 GTCCCTACAAAATCCGGCAGACGACGCTGCGCAAACTGGTCGAGGCGGATACAGCTT 1078
Db 310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCTGACCAACTCGCTACAGCGACCGACGCTTCCGCGCTCCATTCGCGGTACGTC 1138
Db 330 AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValAlaHisAlaGlyTyrAla 349
QY 1139 AAATACCGAAACCGCTCTCAAGCCGCGATCAAACTCTACGAGCTGCAACCCCAACCAT 1198
Db 350 ArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCCGTCCCGCCACA--AAAGACAAAGCGCTGACCGCGCTCCGTAACGACGCTGAT 1255
Db 370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCCAAAACCTTCATTGTGCGGCAAAACGCTATCTTCATCGGCTCATTTCAACCTCGACCC 1315
Db 390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 CGTTCGCGACGCGCTCAATACCGAAATGGCGTGTGTCATCGAAAGCCCAAAATCCAGNA 1375
Db 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
```



```
QY 1490 GAACCCGAGCAAACTTTGGAAACGATCGCGCGGCAAAATCCTATCCTGCTGCTGCCATC 1549
Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGGTTTATTA 1561
Db 490 GluTrpLeuLeu 493

RESULT 11
Q57QL0 SALCH
ID Q57QL0 SALCH PRELIMINARY; PRT; 528 AA.
AC Q57QL0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Putative phospholipase.
GN Name=yndC; OrderedLocustNames=SC1095;
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL; AE017220; AX65001.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 528 AA; 60029 MW; D5175BE93D71CBDB CRC64;
```

Alignment Scores:

```
Score: 4.17e-53 Length: 528
Percent Similarity: 1098.00 Matches: 235
Best Local Similarity: 60.7% Conservative: 88
Query Match: 44.2% Mismatches: 175
DB: 38.5% Indels: 34
Gaps: 8
```

US-10-665-990A-13 (1-1561) x Q57QL0_SALCH (1-528)

```
QY 35 CCGATCAAAACACGACG-----CTCATTTCCCTTTTATGCTCCTT 76
Db 8 ProLeuLysSerProSerGlyArgPheGlnIleLeuLeuProValThrLeuCysLeuAsn 27
QY 77 CTCTGTTTCATGTTCTTCATGGTTGCCCCCCTGGAGAACGACGGAAGCCGTATTTC 136
Db 28 ArgTyrThrLeuSerValMetMetLysLysLeuProGlyPheThrGlnAsp-----Tyr 45
QY 137 AATACTTCAAAACCTGCTCCTCGACACATCCTGCAATCCTGCAATCCGGCACACC----- 187
Db 46 LeuLeuSerLysAlaThrThrLeuProAspLysThrArgLeuGluArgAlaValGluPro 65
QY 188 -----CCTCATACAAACGGGTATCCGACATCTACCTGCTCGACGACCCCGCAAGCC 241
Db 66 LeuCysAlaArgHisProGlyGluCysGlyIleLeuAlaLeuAspAsnSerLeuAspAla 85
QY 242 CTTGCGCGCGCGCCCTTATCGAATCTGCGAACACAGCGCTCGATTGCAATACTAC 301
Db 86 PheAlaAlaArgTyrArgLeuThrGluThrAlaAlaArgThrLeuAspValGlnTyrTyr 105
QY 302 ATTTGGCGCAACGACATTTCCCGGAGCGTGTGTTCAACCTCATGTACCTTCCCGCAGAA 361
Db 106 IleTrpGluAspMetSerGlyArgLeuLeuPheSerValLeuLeuSerAlaAlaLys 125
QY 362 CGCGCGTGGCTGCTGCTGTTGGACGACACACACCGCGGGGTGGACGATCTC 421
Db 126 ArgGlyValHisValArgLeuLeuAspAspAsnThrProGlyLeuAspThr 145

422 CTGCTCGCCCTCGACAGCCATCCCAATATGAGTGGCGCTGTTCACACCCCTCGTCCTCA 481
Db 146 LeuArgLeuLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheSerPhe 165
QY 482 CGCAAAATGGCGCGACCTCGGCTAGCTGACCCAGCTTCCCGCCCTCAACCGCGCATGAC 541
Db 166 ArgThrLeuArgAlaLeuGlyTyrLeuThrAspPheAlaArgLeuAsnArgMethHis 185
QY 542 AACAAATCTTTACCGCGCAACACCGCGCCACCATCTACTCGCGGACGCAATATCGCGAC 601
Db 186 AsnLysSerTyrThrAlaAspGlyValValThrLeuValGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAGTCGGTAGGACACCGCTTTCCCGACCTGGACATCCTCGCCACCGCG 661
Db 206 AlaTyrPheGlyAlaGlyGluProLeuPheSerAspLeuAspValMetAlaIleGly 225
QY 662 AGCGTCGTCGGCGAAGTATCGACACTTCCACCGCTACTCGGCAAGCATTCGCGCCAC 721
Db 226 ProValValAsnAspValAlaAsnAspPheGluArgTyrTrpArgCysSerSerValSer 245
QY 722 AACGCCAGCGCATCATCCGCGGCAACATCGGCAAGGTCTTCAAGCACTC----- 775
Db 246 ThrLeuGlnGlnValLeuSerLeuSerGlu-----GlnGluLeuThrGln 260
QY 776 -----GGATACAAACGACGAGCATCCAGACACGCGCTCCTCGCG 814
Db 261 ArgIleGluLeuProGluSerTrpTyrAsnAspGluIleThrArg-----Arg 276
QY 815 TACCGCGAAACCGTCGACAGCTCGCCCTCTACCAAAATAATACAGCGGACGATCGAC 874
Db 277 TyrLeuHisLysLeuGluThrSerGlnPheMetAlaAspLeuAspArgGlyArgLeuPro 296
QY 875 TGGCAGACGCTCAAAACCGCGCTGATCGACACACCCCTGCAAAAGGACTCGACCGCAC 934
Db 297 LeuIleTrpAlaLysThrArgLeuLeuSerAspProSerLysGlyGluLysAla 316
QY 935 CGCGCAAAACCGCGATTGCGGGAGGCTCAAGACGCGCTCAAAACGCGCAAAAGC 994
Db 317 GlnArgHisSerLeuLeuProGlnArgLeuPheAspValMetGlySerProThrGluArg 336
QY 995 GTCTATCTGGTTTCCCTATTTCCTCCTACAAATCCGCGACAGCGCATGCGCAAAA 1054
Db 337 IleAspIleIleSerAlaTyrPheValProThrArgAlaGlyValAlaGlnLeuLeuAsn 356
QY 1055 CTGCTGACGAGGACGATAGAGTTACCGTCTCTGCAACTCGCTACAGCGGACGACGCTT 1114
Db 357 LeuValArgLysGlyValLysIleAlaIleLeuThrAsnSerLeuAlaAlaAsnVal 376
QY 1115 GCGCGCTCCATTCCGGCTAGCTCAATACCGAAACCGCTGCTCAAAACGCGCATCAAA 1174
Db 377 AlaValValHisAlaGlyTyrAlaArgTrpArgLysLysLeuLeuArgTyrGlyValGlu 396
QY 1175 CTCTACGAGCTGCAACCCCAACCATGCC---GTCCCGCGCCACAAAGACAAAGCGCTGACC 1231
Db 397 LeuTyrGluLeuLysProThrArgGluHisGluThrAlaValHisAspArgGlyLeuThr 416
QY 1232 GGCAGCTCCGTAAACGAGCTCGATGCCAAACCTTCATTGTGGACGGCAACGCGATCTTC 1291
Db 417 GlyAsnSerGlySerSerLeuHisAlaLysThrPheSerIleAspGlySerLysValPhe 436
QY 1292 ATCGGCTCATTTCAACCTCGACCCCGCTCCGACCGGCTCAATACCGAAATGGCGCTGTC 1351
Db 437 IleGlySerLeuAsnPheAspProArgSerThrLeuLeuAsnThrGluMetGlyPheVal 456
QY 1352 ATCGAAACCCCAAAATCGCAGAACAGATGAGGCGCACCCCTCGCGATACACACCCGAA 1411
Db 457 IleGluSerGluThrLeuAlaThrLeuIleHisLysHisPheThrGlnSerGlnArgAsp 476
QY 1412 TACGCTACCGGTTACCTCGACAAACACACCGCCTGCAATGGCACGAT-----CCC 1465
Db 477 AlaIleTrpGlnLeuArgLeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGln 496
QY 1466 GCCACCCGAAAAACCTACCCGAAACGAAACCCGAAAGCAAACTTTTGAAACGCGATCCCGCA 1525
```

```
Db 497 GluGluGluLysValLeuLysGluProAlaThrArgPheTrpGlnArgValLeuVal 516
QY 1526 AAAATCCCTATCCCTGCTCCCATCGAAGGTTTATTA 1561
Db 517 ArgLeuAlaAlaLeuProValGluTrpLeuLeu 528
RESULT 12
ID QBX917_ECO57 PRELIMINARY; PRT; 493 AA.
AC QBX917;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative synthase.
GN Name=yndC; OrderedLocusNames=z1680;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
DR EMBL; AE005174; AAG55792.1; -, Genomic_DNA.
DR PIR; D85666; D85666.
DR GO; GO:0003824; P: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 493 AA; 55981 MW; DCC21D8AAE643457 CRC64;

Alignment Scores:
Pred. No.: 7.38e-53 Length: 493
Score: 1093.50 Matches: 219
Percent Similarity: 65.3% Conservative: 84
Best Local Similarity: 47.2% Mismatches: 144
Query Match: 38.3% Indels: 17
DB: 2 Gaps: 5

US-10-665-990a-13 (1-1561) x QBX917_ECO57 (1-493)
QY 200 GGGGTATCCGACATCTACCTGCTCGACGACCCCGAAGCCCTTGCCGCGCGCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaPheAlaAArgTyrArg 56
QY 260 CTTATCAATCTGCCGACACAGCTCGATTGCAATCTACATTGGCGCAACGACATT 319
Db 57 LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrTrpGlnAspMet 76
QY 320 TCCGGCAGGCTGTGTTCAACCTCATCTACCTGCTCCGCGACAGCGCGGTACGC 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaAlaLysArgGlyValArg 96
QY 380 CTGCTGTGGACGACAAACACGCGGGTGGACGATCTCTGCTCGCCCTCGACGC 439
Db 97 LeuLeuLeuAspAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCGAAGTGGCGCTGTTCAACCCCTCTGCTCTACGAAATGGCGGCACTC 499
Db 117 HisProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
```

```
QY 1550 GNAGGTTTATTA 1561
Db 490 GluTrpLeuLeu 493

RESULT 13
Q9CKM5_PASMU PRELIMINARY; PRT; 514 AA.
ID Q9CKM5_PASMU PRELIMINARY; PRT; 514 AA.
AC Q9CKM5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PM1586.
GN OrderedLocustNames=PM1586;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pm70; PubMed=11248100; DOI=10.1073/pnas.051634598;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
DR EMBL; AE006196; AAK03670.1; -: Genomic DNA.
DR GO; GO:0046740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 514 AA; 57882 MW; C65B0EB65D80F216 CRC64;

Alignment Scores:
Pred. No.: 8 43e-53 Length: 514
Score: 1092.50 Matches: 238
Percent Similarity: 61.7% Conservative: 78
Best Local Similarity: 46.5% Mismatches: 173
Query Match: 38.3% Indels: 23
DB: 2 Gaps: 8

US-10-665-990A-13 (1-1561) x Q9CKM5_PASMU (1-514)
QY 53 CTCATTCCCTTTTANGCTCTCTCTGTTCAATGTTCTTANGTTGCCCCCATGTGAA 112
Db 17 LeuPheSerLeuValIleLeuAlaLeuIleSerTyrGlnArgLeu-----ProThrSer 34
QY 113 GAACGGACGGAAGCGTCATTTCATCTCAATCTTCAACCTGCTCTCGGACACATCTGT 172
Db 35 AspArgProIleSerThrHis-----ProProLeuAsnProAsnGlyLeu 49
QY 173 CAAATCCCGCAC-----ACCCCTCATAAACAGCGCTATCCGACATCTAC 217
Db 50 LeuAlaArgHisIleLeuProGlnValAlaGlnHisProAsn---LeuThrGlyLeuTyr 68
QY 218 CTGCTCGACAGCCCCCAGAACCTTCCGCGCCGCGCGCCCTTATCGAATCTCCGAA 277
Db 69 ProLeuGlyAspGlyLysAlaPheLeuAlaArgLeuAlaLeuSerGluHisAlaGlu 88
QY 278 CACAGCTGATTTCGAATCTACATTTGGCGCAACGACATTTCCGGGAGGCTGTCTTC 337
Db 89 HisThrLeuAspLeuGlnTyrTrpIleTrpHisAsnAspValSerGlyHisLeuLeuLeu 108
QY 338 AACCTCATCTCTTCCGCGAGACCGCGTGCAGCTGCTCTGTTGGACACAC 397
Db 109 GlnSerLeuTyrLysAlaAlaValArgGlyValLysValArgLeuLeuLeuAspAsn 128
QY 398 AACACGCGGGTTGGAGATCTCTGCTCGCCCTCGACAGCCATCCCAATTCGAAGTG 457
Db 129 AsnThrLysGlyMetAspThrIleLeuAlaSerLeuAlaHisProAsnIleGlnIle 148
QY 458 CGCCTGTTCAACCCCTTCGTCTACGCAATGGCGGCACCTCGGCTACCTGACCGACTTC 517
```

```
Db 149 ArgLeuPheAsnProPheMetGlnArgGlnTyrArgTrpLeuGlyPheLeuSerAspPhe 168
QY 518 CCCGCTCAACCCGCGCATGCACAAATTCCTTTTACCGCGCACAAACCGCGCCACCATTA 577
Db 169 PheArgLeuAsnArgMetHisAsnLysSerPheThrAlaAspGlyValMetSerIle 188
QY 578 CTCGGCGACCAATATCGCGGACGAATATCTTAAAGTCGGTGGAGGACACCGTTTTCGCC 637
Db 189 LeuGlyGlyArgAsnIleGlyAspGluTyrPheAspValGlyAsnGlyValLeuPheAla 208
QY 638 GACCTGGACATCTCGCCACCGGAGCTGCTCGCGAAGTATCGCAGACTTCGACCGC 697
Db 209 AspLeuAspValAlaIleThrGlyAlaValThrHisIleGlnThrAspPheAspArg 228
QY 698 TACTGGGCAACCCATTCGCCCCACACCGCCGATCATCCGAGCGGCAACATCGGC 757
Db 229 TyrTrpAsnSerProSerSerTyrProLeuGluSerIleIleIleArgAspProIleThr 248
QY 758 AAGGCTCTTCAAGCACTCGGATACACGACGAAATCCACACACGCGCTCTGGCTAC 817
Db 249 ProPheAsnProLeuProAlaLeuAspAspGluThrGln-----ThrTyr 263
QY 818 CGCGAAACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGCGGAGCATCGACTGG 877
Db 264 LeuLysGlnLeuThrGluLeuProPheAlaLysSerLeuLysAlaGlyThrLeuAlaPhe 283
QY 878 CAGAGCGTCCAAACCCGCTGATCAGCAGACACCCCTCGAAAGGACTCGACCGGACCGC 937
Db 284 ThrTrpAlaGluAlaGluLeuIleSerAspAspProLysLysAlaLeuGlyLysSerLeu 303
QY 938 CGAAACCGCGATTGCGGGAGGCTGCAAGCGGCTCAACAGCCCGCAAAAGCGCTC 997
Db 304 IleGlnAspSerValLeuAlaHisIleAlaProThrMetLeuAsnAlaLysAsnLeu 323
QY 998 TATCTGGTTTCAACCTATTTCCTCCCTACAAAAATCCGGCACAGCGACTGGCAAACTG 1057
Db 324 IleIleValSerProTyrPheValProThrHisValGlyValAspPheLeuSerArgile 343
QY 1058 GTGACGAGCGCATAGACGTTACCTGCTGACCAACTCGCTACAGGCGACCGAGTTGCC 1117
Db 344 SerGlnThrGlyThrGlnValSerIleLeuThrAsnSerLeuGluAlaThrAspValSer 363
QY 1118 CCGCTCATTTCCGCTACGTCAATACCGAAACCGCTGCTCAAGCGGCGATCAAACTC 1177
Db 364 IleValHisSerGlyTyrAlaLysHisArgLysThrLeuLeuGlnLysGlnIleGlnLeu 383
QY 1178 TACGAGCTGCAACCAACCATCGCTGCCGCCCAAAAAAGCAAAAGGC-----CTGACC 1231
Db 384 TyrGluLeuLysPro---HisAlaThrIleGlnMetGluSerSerGlyHisLeuLeuLys 402
QY 1232 GGCAGCTCCGTAAACAGCTCGCATGCCAAAACCTTCATTGTGTGGACGCAACGATCTTC 1291
Db 403 GlyAlaSerSerAlaSerLeuHisAlaLysThrPheThrLeuAspAsnArgTyrLeuPhe 422
QY 1292 ATCGCTCATTTCAACCTCGACCCCGTCCCGAGCTCAATACCGAATGGCGTCTGTC 1351
Db 423 ValGlySerPheAsnMetAspProArgSerAlaMetLeuAsnThrGluMetGlyLeuLeu 442
QY 1352 ATCGAAAGCCCAAAATCGCAGAAACAGATGGAGGCGCACCTCGCGCATACACACCCGAA 1411
Db 443 IleAspSerProGluLeuAlaArgLeuLeuSerAspGlyLeuGlnGlnAsnGlnAlaAsn 462
QY 1412 TAGCGCTACCGGTTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCGCCACC 1471
Db 463 TyrAlaPheSerValLysLeuAsnGluAlaGlnAlaLeuTyrTrpGluThrGlnGluAsn 482
QY 1472 CGAAAA-----ACCTACCCGACGACCCCGAAGCCAAACTTTGGAACGATCCCGCA 1525
Db 483 GlyLysTrpIleArgTyrGluAsnGluProHisThrSerTrpPheLysArgPheSerVal 502
QY 1526 AAAATCTCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
```

Db 503 TrpCysLeuSerTrpLeuProValGluHisLeuLeu 514
 RESULT 14
 Q8ZQ28 SALTY
 ID Q8ZQ28 SALTY PRELIMINARY; PRT: 494 AA.
 AC Q8ZQ28;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Putative phospholipase.
 GN Name=ymdc; OrderedLocusNames=STM1148;
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OC NCBI_TaxID=602;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=LT2;
 RA MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 DR EMBL; AE008749; AAL20078.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR001736; PLD.
 DR Pfam; PF00614; PLDC; 2.
 DR SMART; SM00155; PLDC; 2.
 DR PROSITE; PS00035; PLD; 2.
 KW complete proteome.
 SQ SEQUENCE 494 AA; 56217 MW; 571C01EFC4FCB288 CRC64;
 Alignment Scores:
 Pred. No.: 1.24e-52 Length: 494
 Score: 1089.50 Matches: 221
 Percent Similarity: 63.9% Conservative: 77
 Best Local Similarity: 47.4% Mismatches: 147
 Query Match: 38.2% Indels: 21
 DB: 2 Gaps: 5
 US-10-665-990A-13 (1-1561) x Q8ZQ28_SALTY (1-494)
 QY 200 GGGCTATCCGACATCTACCTGCTCGACGACCCCGAAGCCCTTGGCGCCGCGCGCC 259
 Db 38 GlyGluCysGlyIleLeuAlaLeuAspAsnSerLeuAspAlaPheAlaAlaArgTyrArg 57
 QY 260 CTTATCGAATCTGGCAACACAGCTCGATTGTAATCTACATTTGGCGGCAACAGCAT 319
 Db 58 LeuThrGluMetAlaAlaArgTyrLeuAspValGlnTyrIleTpgIuAspMet 77
 QY 320 TCCGGCAGGCTGCTGTTCAACCTCATGTACCTTCGCGCAGAACGCGGCTGCGGTACGC 379
 Db 78 SerGlyArgLeuLeuPheSerValLeuLeuSerAlaAlaLysArgGlyValHisValArg 97
 QY 380 CTGCTGTTGGACGACAAACACACGCGGGTGGACGATCTCTGCTCGCCCTCGACAGC 439
 Db 98 LeuLeuAspAspAsnThrProGlyLeuAspAspThrLeuArgLeuLeuAspSer 117
 QY 440 CATCCCAATATCGAAGTGGCGCTTTCAACCCCTTCTGCTACGCAATATGGCGGCATC 499
 Db 118 HisProAsnIleGluValArgLeuPheAsnProPheSerPheArgThrLeuAlaGluLeu 137
 QY 500 GGCTACTGACCGACTTCCCGCGCTCAACCGCGCATGACACAAATCTTTACCGCC 559
 Db 138 GlyTyrLeuThrAspPheAlaArgLeuAsnArgMetHisAsnLysSerTyrThrAla 157

RESULT 15

QY 560 GACAAACCGCCACCATCTACGCGGACGCAATATCGCGACGCAATACTTCAAAGTCGCT 619
 Db 158 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 177
 QY 620 GAGGACACCGTTTTCGCGACCTCGACATCTCTGCCACCGGACGCGTCTGCGGAACTA 679
 Db 178 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValAsnAspVal 197
 QY 680 TCGCACACTTCGACCGCTACTGGCAAGCATTCGCGCCCAACAGCCGCGCATCATC 739
 Db 198 AlaAsnAspPheGluArgTyrTrpArgCysSerValSerThrLeuGlnGlnValLeu 217
 QY 740 CGCAGCGCAACATCGGCAAGGCTTCAAGCACTC-----GlnGluLeuThrGlnArgIleGluLeuProGlu 232
 Db 218 SerLeuSerGlu-----GlnGluLeuThrGlnArgIleGluLeuProGlu 232
 QY 776 ---GGATACAGAGCAAAATCATCAGACACGCGCTCTGCGCTACCGGAAACCGTCAA 832
 Db 233 SerTrpTyrAsnAspGluIleThrArg-----ArgTyrLeuHisLysLeuGlu 248
 QY 833 CAGTCGCGCTCTACCAAAAATAACACACGGGACGATCGACTGGCAGAGCGTCAAACC 892
 Db 249 ThrSerGlnPheMetAlaAspLeuAspArgGlyThrLeuProLeuIleTrpAlaLysThr 268
 QY 893 CGCCTGATAGCAGACACCCCTGCAAAAAGGACTCGACCGGACCGCCGCAAAACCGCCGATT 952
 Db 269 ArgLeuLeuSerAspAspProSerLysGlyGlyLysAlaGlnArgHisSerLeuLeu 288
 QY 953 GCGGGAGGCTGCAAGACGCGCTCAACACGCGGAAACGCTCTATCTGTTTCAACC 1012
 Db 289 ProGlnArgLeuPheAspValMetGlySerProThrGluArgIleAspIleSerAla 308
 QY 1013 TATTTCGCTCCTACAAAATCCGACACAGCGCTGCAAACTGGTGGAGGCGGATATA 1072
 Db 309 TyrPheValProThrArgAlaGlyValAlaGlnLeuLeuAsnLeuValArgLysGlyVal 328
 QY 1073 GAGCTTACCCTCTGACCAACTCGCTACGCGGACCGGCTGCGCGCTCATTCGCGC 1132
 Db 329 LysIleAlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGly 348
 QY 1133 TAGCTCAATACCGAAACCGCTGCTCAAGCGGCACTCAAACTCTACGAGTCAACCC 1192
 Db 349 TyrAlaArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysPro 368
 QY 1193 AACCATGCC---GTCCCGCCACAAAAGAGCAAGGCTGACCGGACGCTCGTAAACGAC 1249
 Db 369 ThrArgGluHisGluThrAlaValHisAspArgGlyLeuThrGlyAsnSerGlySerSer 388
 QY 1250 CTGCATGCCAAAACCTTCATTGTGGACGGCAACGCACTTTCATCGGCTCATTCACACTC 1309
 Db 389 LeuHisAlaLysThrPheSerIleAspGlySerLysValPheIleGlySerLeuAsnPhe 408
 QY 1310 GACCCCGTTCCGACGCGCTCAATACGAAATGGGCTCGCTCATCGAAAGCCCAAAATC 1369
 Db 409 AspProArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeu 428
 QY 1370 GCAGAACAGATGAGCGCACCTCTCGCGATACACACCGCAATACGCTACCGGTTACC 1429
 Db 429 AlaThrLeuIleHisLysArgPheThrGlnSerGlnArgAspAlaIleTrpGlnLeuArg 448
 QY 1430 CTCGACAAACACAAACCGCTGCAATGGCACAT-----CCCGCCACCCGAAACACCTAC 1483
 Db 449 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGlnGluGluLysValLeu 468
 QY 1484 CCGAACGACCCGAAAGCCAACTTTGGAAACGATCGCGCAAAATCTTATCCCTGCTG 1543
 Db 469 LysLysGluProAlaThrArgPheTrpGlnArgValLeuValArgLeuAlaIleLeu 488
 QY 1544 CCCATCGAAGTTTATTA 1561
 Db 489 ProValGluTrpLeuLeu 494


```

Db 450 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGlnGluGluValLeu 469
QY 1484 CCGAACGAAACCGGAAACCTTTGGAAAGCGATCGCGCAAAATCCTATCCCTG 1543
Db 470 LysLysGluProAlaThrArgPheTrpGlnArgValLeuValArgLeuAlaAlaLeu 489
QY 1544 CCCATCGAAGTTTATTA 1561
Db 490 ProValGluTrpLeuLeu 495

RESULT 16
QSPGY9_SALPA
ID QSPGY9_SALPA PRELIMINARY; PRT; 495 AA.
AC QSPGY9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein ymdC.
GN Name=ymdC; OrderedLocuNames=SPA1703;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=1531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274 (2004).
DR EMBL; CP000026; AAU77627.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 56310 MW; 8BAADE58874983A2 CRC64;

Alignment Scores:
Pred. No.: 1.24e-52 Length: 495
Score: 1089.50 Matches: 221
Percent Similarity: 63.9% Conservativeness: 77
Best Local Similarity: 47.4% Mismatches: 147
Query Match: 38.2% Indels: 21
DB: 2 Gaps: 5

US-10-665-990A-13 (1-1561) x QSPGY9_SALPA (1-495)
QY 200 GGGCTATCCGACATCTACCTGTCGACGACCCCAAGACCCCTTGCCGCCGCGGCC 259
Db 39 GlyGluCysGlyIleLeuAlaLeuAspAsnSerLeuAlaPheAlaAlaArgTyrArg 58
QY 260 CTTATCGAATCTCGGAACACAGCTCGATTTCGAATCTACATTTGGCCGCAACGACATT 319
Db 59 LeuThrGluMetAlaAlaArgThrLeuAspValGlnTyrTrpIleTrpGluAspMet 78
QY 320 TCCGGCAGGCTGTGTTCAACCTCATGTACTTCCCGCAGAACCGCGCGTCCGATACGC 379
Db 79 SerGlyArgLeuLeuPheSerValLeuSerAlaAlaLysArgGlyValHisValArg 98
QY 380 CTGCTGTTGACGACAAACAACGCGCGGTTGGACGATCTCTGCTCGCCCTCGACAGC 439

```

```

Db 99 LeuLeuLeuAspAspAsnAsnThrProGlyLeuAspAspThrLeuArgLeuLeuAspSer 118
QY 440 CATCCCAATATCGAAGTCGCGCTGTTCAACCCCTTCTCTACGCAATGCGCGCATC 499
Db 119 HisProAsnIleGluValArgLeuPheAsnProPheSerPheArgThrLeuArgAlaLeu 138
QY 500 GGGTACCTGACCGACTTCCCGCCCTCAACCGCGCATGCAACAATACTCTTTACGCC 559
Db 139 GlyTyrLeuThrAspPheAlaArgLeuAsnArgArgMethHisAsnLysSerTyrThrAla 158
QY 560 GACAAACCGCGCCACCATACTCGCGGACGCAATATCGCGACGCAATACTTCAAGTCGT 619
Db 159 AspGlyValValThrLeuValGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 178
QY 620 GAGGACACCGCTTTTCGCGACCTGGACATCTCGCCACCGCGACGCTGTCGGCAAGTA 679
Db 179 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValAlaAsnAspVal 198
QY 680 TCGCAGCAGCTTCGACCGCTACTCGGCAAGCATTTCCGCCCAACGCGCCGCGCATCATC 739
Db 199 AlaAsnAspPheGluArgTyrTrpArgCysSerValSerThrLeuGlnGlnValLeu 218
QY 740 CGCAGCGGCAACATCGGCAAGGCTCTTCAAGCACTC----- 775
Db 219 SerLeuSerGlu-----GlnGluLeuThrGlnArgIleGluLeuProGlu 233
QY 776 ---GGATACAAACGACGAAACATCCAGACACGCGCTCTCGCTACCGCAACCGTCGAA 832
Db 234 SerTrpTyrAsnAspGluIleThrArg-----ArgTyrLeuHisLysLeuGlu 249
QY 833 CAGTCGCGCCCTTACCAAAAAATACAGACGGGACGATCGACTGGCAGAGCGCTCCAAACC 892
Db 250 ThrSerGlnPheMetAlaAspLeuAspCysGlyThrLeuProLeuIleTrpAlaLysThr 269
QY 893 CGCTGATCAGCAGCACCCCTCGCAAAAGGACTCGACCGACCGCCGCAACCGCGGATT 952
Db 270 ArgLeuLeuSerAspAspProSerLysGlyGluGlyAlaGlnArgHisSerLeuLeu 289
QY 953 GCCGGGAGGCTCAAGACGCGCTCAACACAGCCCGAAAAAGCGCTCTATCTGGTTTCAACC 1012
Db 290 ProGlnArgLeuPheAspValMetGlySerProThrGluArgIleAspIleSerAla 309
QY 1013 TATTTCGCTCTACAAAAATCCCGCACAGACGCTCGCAAAACTGGTGACGAGCGGCATA 1072
Db 310 TyrPheValProThrArgAlaGlyValAlaGlnLeuLeuAsnLeuValArgLysGlyVal 329
QY 1073 GAGCTTACGCTCTGACCACTCGCTACAGGCGACCGAGTTCGCCGCGTCCATTCCGGC 1132
Db 330 LysIleAlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGly 349
QY 1133 TAGCTCAATACCGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCC 1192
Db 350 TyrAlaArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysPro 369
QY 1193 AACCATGCC---GTCCCGCCCAAAAGACAAAGGCTGACCGCGAGCTCCGTAAACCAGC 1249
Db 370 ThrArgGluHisGluThrAlaValHisAspArgGlyLeuThrGlyAsnSerGlySerSer 389
QY 1250 CTGCATGCCAAACCTTCATTGTGGACGGCAACGCATCTTTCATCGGCTCATTTCAACCTC 1309
Db 390 LeuHisAlaLysThrPheSerIleAspGlySerLysValPheIleGlySerLeuAsnPhe 409
QY 1310 GACCCCGTTCGCGACGCGCTCAATACGAAATGGCGTCTCATCCAAAGCCCAAAATC 1369
Db 410 AspProArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeu 429
QY 1370 GCAGAACAGATGAGCGCACCTCGCCGATACACACCCCGAATACCGCTACCGCGTACC 1429
Db 430 AlaThrLeuIleHisLysArgPheThrGlnSerGlnArgAspAlaAlaTrpGlnLeuArg 449
QY 1430 CTCGACAAACACAAACCGCTCGAATGGCACGAT-----CCGCGCACCGCAAAACCTAC 1483
Db 450 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGlnGluGluLysValLeu 469

```

Qy	38	ATGAAACACGACGCTCATTTCCCTTTTATGCTCTCTCTGTTCAATGTTCTTTCATGG	97
Db	1	MetLeuThrLeuSerLeuValThrValValGlyIleAsnGlyIleCysSerThrLeuProLys	20
Qy	98	TTGCCCCCACTGGGAAGAACGCGAAGCGCTCATTTCAATATCTCCAAACCTGTCCTC	157
Db	21	HisGlyPro---GluValThrGlnTyAlaTyAspIleAspThrSerGlnThrSerLeu	39
Qy	158	CTGGACAAACATCTTGCAATCCGGCACACCCCTCATAAACAGGGCTATCCGACATCTAC	217
Db	40	AlaGlnIleIleThrProLeuLysGln-----GlnAsnLeuGlyLeuThrGlyTyHis	57
Qy	218	CTGCTCGACAGCCCCACGAAGCCCTTGGCCGCCCGCGCCCTTATCGAATCTGCCGAA	277
Db	58	MetLeuTyGluProLeuGluAlaLeuAlaAlaArgLeuArgLeuIleAspLysAlaGlu	77
Qy	278	CACAGCCTCGAATTTGCAATACATATTGGGCGCAACGACATTTCCGGCAGGCTGCTGTC	337
Db	78	LysThrLeuAspLeuGlnIleTyIleThrAspAsnAspLysIleGlyAlaLeuAlaLeu	97
Qy	338	AACCTCATGTACCTTGGCCGACAGACGGCGCTGGCGGTACGCCCTGCTTGGAGACCAAC	397

```

Db      438 AlaTyrlEuAnThrGluIleGlyValIleLeuAsnSerProProLeuAlaMetaVal 457
Qy      1382 GAGCGCACCTTCGCGCATACACACACCCGAATACGCCTACCGCGTTACCTCGACAAACAC 1441
Db      458 HisGlnThrMetAspGlnGlnLeuThrHisTyrlaTyrlLysLeuValLeuAspAlaAsn 477
Qy      1442 AACCGCTGCATGGCAGATCCCGCCACC-----CGAAAAAAGCTACCCGGAACGAAACC 1495
Db      478 HisLysIleAsnTrpTyArgGlnThrSerThrGlyThrLysIleTyThrLysGluPro 497
Qy      1496 GAAGCCAAATCTTGGAACGATCCGCGCAAAATCCTATCCCTGCTGCCCATCGAAGGT 1555
Db      498 ArgMetLysTrpTrpGlnLysAlaGlyIleLysLeuIleSerTrpLeuProIleGluGly 517
Qy      1556 TTAATTA 1561
Db      518 PheMet 519

RESULT 18
Q4LR64_9BURK
ID      Q4LR64_9BURK PRELIMINARY; PRT; 516 AA.
AC      Q4LR64;
DT      13-SEP-2005 (TrEMBLrel. 31, Created)
DT      13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Phospholipase D/Transphosphatidylase precursor.
GN      ORFNames=Bcen2424DRAFT_3266;
OS      Burkholderia cenocepacia HI2424.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC      Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX      NCBI_TaxID=331212;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=HI2424;
RG      US DOE Joint Genome Institute (JGI-PGF);
RA      Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA      Hammon N., Israni S., Pitluck S., Richardson P.;
RT      "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT      HI2424."
RL      Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=HI2424;
RG      US DOE Joint Genome Institute (JGI-ORNL);
RA      Larimer F., Land M.;
RT      "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT      HI2424."
RL      Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC      -1- CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
KW      EMBL; AAHL01000025; EAM18523.1; -; Genomic_DNA.
DR      Signal.
FT      SIGNAL.
SQ      SEQUENCE 516 AA; 57599 MW; 56B9D87A8D18A60E CRC64;

Alignment Scores:
Pred. No.: 5,08e-41 Length: 516
Score: 881.50 Matches: 214
Percent Similarity: 55.0% Conservative: 77
Best Local Similarity: 40.5% Mismatches: 203
Query Match: 30.9% Indels: 35
DB: 2 Gaps: 11

US-10-665-990A-13 (1-1561) x Q4LR64_9BURK (1-516)
Qy      38 ATGAAACACGCGACCTCATTTCCCTTTTATGCTCTCTCTGTTCATGTTCTTCATGG 97
Db      2 IleArgLeuArgSerTrpGlyAlaIleLeuSerLeuMetLeuAlaCysAlaSer 21
Qy      98 TTGCCCCACTGGAGAACCGACGGAAGCGGTCAATTTCAATCTTCCAAACCTGCCTC 157
Db      22 LeuProGlnAlaAspArgAlaProThrHisAlaPheThrAspThrAspThrArg 41

```

Qy	158	CTGACAAACATCTCTGCAAAATCCGGCACACCCCTCATAAACACGGGCTATCCGACATCTAC	211
Db	42	LeuGlyValAlaPheArgGlnGlnAlaAlaHis---ProGlyGlnAspAlaPheHis	60
Qy	218	CTGCTCAGACACCCCAAGAGCCCTTGCGCCCGCCGCGCCCTTATCGAATCTCCGCA	277
Db	61	LeuLeuThrAspProValAspAlaLeuAspAlaArgValLeuLeuAlaAspArgAlaGlu	80
Qy	278	CACAGCTCTGATTTGCAATACTACATTTGGCGCAACGACATTTCCGGCAGGCTGTCTGTC	337
Db	81	ArgSerLeuAspLeuGlnTyrTyIleTrpHisAspAspLeuThrGlyHisGluLeuAla	100
Qy	338	AACCTCATGTACTTGC CGGAGACGGCGGTGCGCTACGCTCTGTTCTTCGAGCAGCAAC	397
Db	101	AspAlaValMetArgAlaAlaAspArgGlyValArgValArgAlaLeuLeuAspLeu	120
Qy	398	AACACGCGCGGTTGGAGCATCTCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTG	457
Db	121	GlyThrAsnAlaAspAspArgLysLeuLeuGluIleSerAlaHisProAsnIleGluIle	140
Qy	458	CGCTGTTCAACCCCTTCGTCCTACGCAATATGGCGGCACATCGGCTACTCAGCAGCTTC	517
Db	141	ArgLeuPheAsnProValAlaSerArgArgPheLysLysLeuGlyThrLeuPheGluPhe	160
Qy	518	CCCCGCTCAACCGCCGATGCACAAACAAATCTTTTACCGCGCACACCGCGCCACCATTA	577
Db	161	SerArgValAsnArgArgMetHisAsnLysAlaMetIleAlaAspAsnGlnAlaAla	180
Qy	578	CTGCGCGGACGCAATATCGCGCAGCAATACTTCAAAGTCGGTGAGGACACCGTTTTCGCC	637
Db	181	ValGlyGlyArgAsnIleGlyAspGluTyrPheGlyAlaSerSerMetLeuGluPheGly	200
Qy	638	GACTGAGACATCTCGCCACCGCGCAGCGTCTGCGCGAAGTATCGCAGCACTTCGACCGC	697
Db	201	AspLeuAspValValValHisGlyProValValLysAspIleSerThrGluPheAspThr	220
Qy	698	TACTGGCAAGCCATTCGCGCCCAACACCGCCATCATCCGCGCGGCAAGCATCGGC	757
Db	221	PheTrpAsnSerProTyrAlaTyrProValAspThrLeuValGlyHisAspAlaAlaPro	240
Qy	758	AAGGTTCTTCAAGCACTCGGATACACAGCAGAAACATCCAGACAGCGCTCTCGGCTAC	817
Db	241	GlyGlyLeu-----GluArgGluArgGluArgLeuArgAspThr	253
Qy	818	CGCGAAACCCCTGCAACAGTCGCCCTCTAC-----CAAAAATA-----	856
Db	254	LeuArgAlaMetGluAspAsnProTyrValLeuGluAlaArgGlnArgLeuAspArgile	273
Qy	857	-----CAGACGGGACGATCGATGCGCAGAGCGTCCAAACCCGCTGATCAGCGAC	907
Db	274	ValHisGlyGlnGlyThrGluLeuSerTrpGly-----HisAlaThrValLeuTyrAsp	291
Qy	908	ACCCCTGCAAAA-----GGACTCGACCGCGACCGCCGCAACCGCCGATTGCC	955
Db	292	AspProSerLysIleAlaHisProProLysAspArgAspGlyHisLeuMetPro-----	309
Qy	956	GGGAGCGCTCAAGACCGCTCAACACACGCCGAAAAAAGCGTCTATCTGTTTTCACCTAT	1015
Db	310	---GlnLeuArgAlaLeuAlaLeuGlnProArgHisAspLeuIleIleValSerProTyr	328
Qy	1016	TTTCGTCCTTACAAATCCCGCACAGCGCACTGGCAAACTGGTGTCAGGACGGCATAGC	1075
Db	329	PheValProGlyLysGluValValGluArgLeuArgAlaLeuThrAlaArgGlyValArg	348
Qy	1076	GTTACCGCTCTGACCAACTCGCTACAGCGCACCGACCTTGGCGCGCTCCATTTCGGGTAC	1135
Db	349	ValThrValLeuThrAsnSerLeuAlaSerThrAspValAlaAlaValHisAlaGlyTyr	368
Qy	1136	GTCAATATCCGAAAAACCGTCTCAAAGCCGGATCAAATCTTACGAGCTGCAACCCCAAC	1195
Db	369	ArgArgTyrArgSerAspLeuValAsnAlaGlyValArgLeuTyrGluArgTyrArgProSer	388

Qy	1127	TCCGGCTACGTCAAATACCGAAACCCGCTGCTCAAGACCGGCATCAAACTCTACGAGCTG	1183
Db	364	GlyTrpTyrAlaLysTyrArgGluAspLeuLeuGluSerGlyIleLysLeuTrpGluVal	383
Qy	1187	CAACCCCAACCATGCGGTCCCGCCGACAAAAGACAAA---GGCTGACCGGAGCTCCGTA	1243
Db	384	LysSerSerAlaLysLeu-----LysSerLysTrpSerLeuThrGlySerSerArg	400
Qy	1244	ACCAGCTGCATGCTCAAAACCTTCATTTGTGGACGGCAACGCATCTTTCATCGGCTCATTC	1303
Db	401	AlaSerLeuHisAlaLysAlaMetThrIleAspAspLysThrLeuPheValGlySerMet	420
Qy	1304	AACCTCGACCCCGCTTCGCGACGCGCTCAATACCCGAAATGGCGGTGCTCATCGAAAGCCCC	1363
Db	421	AsnTrpAspProArgSerAlaAlaLeuAsnThrGluMetAlaValValIleGluGlnPro	440
Qy	1364	AAATATCGACAGACAGATGGAGCGCACCTTCGGCGATACCAACCCGAAATACGCTACCGC	1423
Db	441	GluTyrValGlnAlaPheLeuAlaLysLeuProSerGlnLeuLysAspAsnAlaTyrArg	460
Qy	1424	GTTATCCCTCGACAAACACACCGCGCTGCAATGGCAGCATCCGCGCACCCGAAACCTAC	1483
Db	461	LeuThrLeu---ArgAspGlyAspIleValTrpThrAsnThrLysThrGlyGluGluTyr	479
Qy	1484	CCGAAACCAACCCGAAGCCAACTTTGGAAACGATCGCCGCAAAATCCCTATCCCTGCTG	1543
Db	480	AspSerGluProGluAlaGlyValPheArgGluGlyAlaTrpPheSerGlyIleLeu	499
Qy	1544	CCCATCGAAGGTTTATTATTA 1561	
Db	500	ProlleGluAspGlnLeu 505	

RESULT 20

Q8D641_VIBVU

ID Q8D641_VIBVU PRELIMINARY; PRT; 501 AA.

AC Q8D641;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Phosphatidylserine/phosphatidylglycerophosphate/ cardiolipin synthase.

DE GN OrderedLocusNames=VV20702;

OS Vibrio vulnificus.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrio.

OX NCBI_TaxID=672;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=CMCP6;

RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H., Choy H.E.;

RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR ENBL; AE016810; AA007638.1; -; Genomic_DNA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.

DR SMART; SM00155; PLDC; 2.

DR PROSITE; PS50035; PLD; 2.

KW Complete proteome.

SQ SEQUENCE 501 AA; 56542 MW; 4A7DA73EB0A836B2 CRC64;

Alignment Scores:	
Pred. No.:	1,61e-40
Score:	872.50
Percent Similarity:	56.6%
Best Local Similarity:	39.1%
Query Match:	30.6%
DB:	2
Length:	501
Matches:	198
Conservative:	89
Mismatches:	211
Indels:	9
Gaps:	6

US-10-665-990A-13 (1-1561) x O8D641 VIEWU (1-501)

QY	47	CGCAGCCTCATTTCCCTTTTATGCGCTCCTTCTCTGTTCTATGTTCTTCTATGTTGTCGCCCA	106
DB	2	ArgValLeuValValAlaSerCysPheGlyLeuLeuAlaCysSerAlaSerAsnProLeu	21
QY	107	TTGGAAGAACCGACGGAAAGCGCTCATTTTCAATCTACTTCCAAA---CCTGCTCCTCTGGAC	163
DB	22	ValThrGluLysLeuSerTyrHisSerGlyTyrGlnGluGlySerThrLeuAlaLys	41
QY	164	AACATCTCTGCAATCCGGCACACCCCTCATACAACAGGGCTATCGACATCTACCTGCTC	223
DB	42	TyrTyrGluGlnLeuArgProGluThrValLysAspAsnLeuThrGlyPheHisSerLeu	61
QY	224	GACGACCCCAAGAGCCCTTGGCCCGCGCGCTTATCGAATCTGCCAATCTGCCAACACACAGC	283
DB	62	AsnAsnGlyAspAspAlaLeuLeuAlaArgLeuAlaLeuIleThrSerAlaGlnHisSer	81
QY	284	CTCGATTTTGCATACTACATTTTGGCGCAACGACATTTCCGGCAGCGTGTGTTTCAACCTC	343
DB	82	LeuAspLeuGlnTyrTyrLeuPheGlyAspAspGluThrSerLysIleIleThrTrpArg	101
QY	344	ATGTACTCTTGGCGAGAACGGCGTGGCTAGCCCTGCTGTTTGGAGACAAACACAGC	403
DB	102	LeuTyrGluAlaAlaGlnArgGlyValLysValArgLeuLeuLeuAspMetGlnArg	121
QY	404	CGCGGTTTGGACGATCTCCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTCGCGCTG	463
DB	122	ArgAsp---AspGlnValMetAlaLeuLeuAsnAspHisProAsnIleGluIleArgLeu	140
QY	464	TTCAACCCCTTCTGCTACGCAAAATGGCGCGCATCTCGCTTACCTGACCGATTCCTCCCGC	523
DB	141	PheAsnProTyrThrAsnArgPhePheArgGlyLeuSerPheLeuThrAspPheSerArg	160
QY	524	CTCAACCCCGCATGCACAAATCTTACCGCCGACACCGCGCCACCATCTCGGC	583
DB	161	LeuAsnHisArgMetHisAsnLysSerMetThrAlaAspGlyValAlaAlaIleValGly	180
QY	584	GGACGCAATATCGCGACGAATACTTCAAAGTCGGTGAGGACACCGCTTTTCCGCCACCTG	643
DB	181	GlyArgAsnIleGlyAsnGluTyrPheSerPheSerThrGlnValGluPheGlyAspPhe	200
QY	644	GACATCTCTCGCACCGGCGAGCTCGTTCGGCGAAGTATCGCACGATCTCGACCGCTACTGG	703
DB	201	AspValLeuLeuPheGlyProAlaValValGluThrAlaThrGlnPheAspLeuTyrTrp	220
QY	704	GCAAGCCATTCGGCCCAACAGCCACCGCGCATCATCCGACGGGCAACATCGCGAAGGT	763
DB	221	AsnSerAspGlnThrGlnSerValSerAlaLeuValSerLeuSerAspTyrGluThrGln	240
QY	764	CTTCAAGCACTCGGATACAAACGACGAAACATCCAGACAGCGCTCTCGCTACCGCGAA	823
DB	241	TyrAlaLeuGluAspLeuValAspValAsnAlaLeuGluAlaProPheArgAspGlyGlu	260
QY	824	---ACCGTTCGAAGTCCGCCCTCTACCAAAAAATACAGACGGACGCGATCGACTGGCAG	880
DB	261	TyrAspIleSerGlnLeuGluLeuPheGluHisLeuLysHisGlyThrLeuLysLeuTyr	280
QY	881	AGCGTCCAAACCCGCTGATCAGCGACACCCCTCGAAAGACATCGACCGGACCGCGC	940
DB	281	TrpGlyGluAlaValLeuTyrTrpAspLeuProAspLysValGluThrArgAspSerGln	300
QY	941	AAACCGCGGATTCGGGAGGCTCGACAGCGCTCAACAGCCGCGCAAAAAAGGCTCTAT	1000
DB	301	-----ValAlaGluGlnLeuAlaGluArgLeuAsnAsnValThrAspSerIleLeu	317
QY	1001	CTGGTTTCAACCTATTTTCCTTACAAAAATCCGGCACAGCACTGGCAAACTGTGTG	1066
DB	318	IleIleSerProTyrPheValProThrGluAlaGlyThrLeuGlyLeuLysArgLeuAla	337
QY	1061	CAGACCGCATAGAGTTTACCGTCTGACCAACTCTCGCTACAGGCGACCGAGTTGCGCC	1120
DB	338	GluSerGlyValLysValThrIleValThrAsnSerLeuAlaSerAsnAspValPheAla	357

```

QY 1121 GTCCATCCGGCTACGTCAAAATACCGAAACCGTCTCTCAAGCGGCATCAAACTCTAC 1180
Db ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
358 ValHisGlyTyrTyrAlaLysTyrArgGluGluLeuLeuGluSerGlyIleThrLeuTrp 377
QY 1181 GAGTGCACCAACCATCGCTCCCGCCACAAAAGCAAAAGGCGCTGACCGGAGCTCC 1240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
378 GluIleLysAlaSerAlaGluIle-----LysArgGlnTrpSerLeuIleGlySerGln 395
QY 1241 GTAAACAGCTGTCATGCCAAAACCTTCATTGTGGACGGCAAAACGATCTTCATCGGCTCA 1300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 ArgAlaSerLeuHisAlaLysValIleValIleAspAlaAspPheIleValGlySer 415
QY 1301 TTCACCTCGACCCCGTCCGACGGCTCAATACCGAAATGGGCGTCTCATCGAAAGC 1360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 MetAsnTrpAspProArgSerAlaTyrIleAsnThrGluMetAlaValHisIleThrHis 435
QY 1361 CCCAAAATCGCAGACAGATGAGCGCACCTCGCGGATACACACCCGGAATACGCCCTAC 1420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
436 ProGluTyrValLysLysAlaThrValGlnLeuAsnSerAlaLeuProLysAspAlaTyr 455
QY 1421 CGCGTTACCTCGCAACACACACCGCTGCAATGGCAGATCCCGCCACCGCAAAACC 1480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
456 TyrValGluIleArgGluGlyAsn---LeuHisTyrThrAspLeuGlyAsnGlyLysVal 474
QY 1481 TACCGAAGAACCGGACGCAAACTTTGGAAACCATCGCGCGCAAAATCCTATCCCTG 1540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 TyrAspSerGluProGluSerSerIleTrpArgSerIleGlyAlaTrpLeuSerGlyVal 494
QY 1541 CTGCCCTCAGAGGTTTATTA 1561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 LeuProIleGluGlyMetLeu 501

```

RESULT 21

```

Q8D861_VIBVU
ID Q8D861_VIBVU PRELIMINARY; PRT; 524 AA.
AC Q8D861
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=VWI3126;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxId=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016807; AAO11447.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00335; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 524 AA; 59929 MW; 8C88CB38C8CADA7 CRC64;

```

Alignment Scores:

```

Pred. No.: 2.38e-40 Length: 524
Score: 869.50 Matches: 209
Percent Similarity: 56.5% Conservative: 88
Best Local Similarity: 39.7% Mismatches: 190
Query Match: 30.5% Indels: 39
DB: 2 Gaps: 14

```

US-10-665-990A-13 (1-1561) x Q8D861_VIBVU (1-524)

QY 29 ATATCCCGCATGAAACA-----CCGAGCCTCATTTCCCTTTTATGCCTCTCTC 79

```

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 LeuGlyProLeuArgThrPheIleHisLysAlaLeuThrTyrCysSerValMetLeuLeu 36
QY 80 TGTTCATGTTCTTTCATGTTGGTCCCGCCCACTGAAACGACGAGAAAGCGGTCAATTCAAT 139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
37 CysAlaCysThrSer-----LeuGluAsnHisSerPro-----PheAsp 49
QY 140 ACTTCCAAACCTGCTCCTCTG-----GACAAATCATCTCGCAA-----ATCCGG 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 LysGlnProSerTyrGlnLeuGlyTyrGlnAlaAspSerArgLeuSerAlaTyrLeuAsn 69
QY 182 CACACCCCTCATAC---AACGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAA 238
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
70 HisHisProGlnAspArgGluAsnLeuThrAlaPhePheProLeuAspLysGlyHisasp 89
QY 239 GCCCTTGGCGCGCGCGCCCTTATCGAATCTCCGAAACACAGACCTCGATTGCAATAC 298
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
90 AlaLeuLeuAlaArgLeuAlaLeuIleGluAlaAlaAspLysThrLeuAspLeuGlnTyr 109
QY 299 TACATTTGGCCCAACGACATTTCCGGCAGGCTGTGTTCACCTCATGTACTTGCCTGCA 358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
110 TyrIlePheArgAsnAspGluThrGlyGlnLeuLeuThrTyrArgLeuPheGluAlaAla 129
QY 359 GAACGCGCGTGGCGCTACGCTGCTGTGGACGACAAACACGCGGGGTTGGACGAT 418
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 GlnArgGlyValArgValArgIleLeuLeuAspMetGlnLysArgAsn---AspGlu 148
QY 419 CTCCTGCTCGCCCTCGACGACCATCCCAATATCAAGTGGCGCTTCAACCCCTTCGTC 478
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 GlyLeuAlaArgLeuSerAlaHisProAsnIleGlnIleArgLeuPheAsnProHisGln 168
QY 479 CTACGCAAAATGGCGGCACTCGGCTACTGACCGACTTCCCGCTCAACCGCGGCAATG 538
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 TyrArgThrAlaArgThrLeuAlaMetAlaSerAspPheSerArgLeuAsnArgArgMet 188
QY 539 CACAACAAATCCTTTTACCGCGCAACCGCGCCACCATCTACGCGGACGCAATATCGGC 598
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 HisAsnLysSerLeuThrAlaAspSerValValAlaIleValGlyArgAsnIleGly 208
QY 599 GACGAATACTTCAAAGTGGTGGAGACACCGCTTTTGGCCGACCTGGACATCCTCGCCACC 658
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 AsnGluTyrPheSerValAsnSerProValGluPheGlyAspPheAspLeuMetLeuTyr 228
QY 659 GGCAGCGTCTCGCGGAGTAGTATCGACGACTTCGACCGCTACTGGGCAAGCATTCGCCCC 718
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
229 GlyAsnSerValGluGlnThrAlaGluGlnPheAspLeuTyrTrpAsnSerLeuHisAla 248
QY 719 -----CACAACGCCACGCGCATCATCCGCGGACGCGCAACATCGGC 757
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 ValProIleGluTyrLeuThrAspAsnProIleProValThrGlu----- 263
QY 758 AAGGCTTCTTAAGCACTCGGATACAACGACGAAACATCC---AGACACGCGCTCTCGGC 814
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 GluGluLeuGlnAla-----TrpLeuLysGluThrGlnLeuGluGlnLysPheThrGln 281
QY 815 TACCGGAAACCGTCAACAGTCCGCCCTCTACCAAAAATAATACAGCGGACGATCGAC 874
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 GlyArgTyrAspPheThrGlnLeuGlnLeuTyrGlnGlnPheThrAspLysSerLeuVal 301
QY 875 TGGCAGAGCGTCCAAACCCCGCTGATCAGCGACACCCCTCGAAAAGAGCATCGACCGCAG 934
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
302 TrpTyrTrpGlyLysGlyValTrpTyrAspLeuProAspLys-----ValAsp 318
QY 935 CGCGCAAAACCGCGATTTCCGGGAGGCTGCAAGACGCGCTCAACACGCCCGCAAAAAGC 994
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 ThrGlnAlaProGlnLeuAlaAspAsnLeuAlaSerLeuLeuArgThrValLysAspSer 338
QY 995 GTCTATCTGTTTACACCTATTTCTCCCTACAAATCCGCGACAGACGACTGCAAAA 1054
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 LeuValLeuIleSerProTyrPheValProThrGluArgGlyThrGlnAlaLeuValGlu 358
QY 1055 CTGTGTCAGGACGCGCATAGAGCTTACCCTCTGACCAACTCGTACAGCGGACCGAGCTT 1114
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 359 AlaAlaHisArgGlyValAspIleThrIleValThrAsnSerLeuAlaSerAsnAspVal 378
QY 1115 GCCGCGCTCCATTCGGCTACGTCAATACCGAAACCGCTGCTCAAGCCGGCATCAA 1174
Db ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 PheAlaValHisGlyTrpTyAlaLysTyrArgGlnAspLeuValGluAlaGlyIleGln 398
QY 1175 CTCACGAGCTGCAACCAACCATGCGCTCCCGCCCAACAAAGACAAAGGCTGACCGGC 1234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 LeuTrpGluThrLysAlaSerAlaArgIleAspSer-----LysTrpSerPheThrGly 416
QY 1235 AGCTCCGTAACACGCGTCATGCCCAACCTTCATTGTGGAGCGCAACGCGATCTTCATC 1294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 SerSerArgSerLeuHisAlaLysValLeuLeuIleAspHisArgLeuLeuPheAla 436
QY 1295 GGCTCAATCAACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGGGTGTGTCATC 1354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
437 GlySerMetAsnTrpAspProArgSerAlaLeuLeuAsnThrGluMetAlaAlaValIle 456
QY 1355 GAAAGCCCCAAATCGCAGAAACAGATGGAGCGCACCTTCGCGGATACCCACCCGAATAC 1414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
457 GluHisProAspTyrValGlnSerSerGluAlaLysLeuProMetGlyLeuGluThrAsn 476
QY 1415 GCCTACCGCTTACCTCGACAAACACACCGCTGCAATGGCAGCATCCCGCCACCCGA 1474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 AlaTyrGlnValArgMet---LysAsnGlyGluValAlaTrpPheAspHisGlnSerGln 495
QY 1475 AAAACCTACCCGGAACGAAACCGAAACCTTTGGAAACGCATCGCGCAAAAATCCTA 1534
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
496 ValTrpPheAspSerGluProGluAlaThrValTrpArgLysIleGlyAlaTrpPheAla 515
QY 1535 TCCTGTGTCCTCCATCGAA 1552
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
516 GlyIleLeuProIleGlu 521

RESULT 22
Q7MMB3_VIBVY
ID Q7MMB3_VIBVY PRELIMINARY; PRT; 554 AA.
AC Q7MMB3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VV1159.
GN OrderedLocusNames=VV1159;
OS Vibrio vulnificus (strain VJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA00037; BAC93923.1; -, Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR PROSITE; PS0035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 554 AA; 63341 MW; 8A7980BB096AB6C8 CRC64;

```

```

Alignment Scores:
Pred. No.: 5.17e-40 Length: 554
Score: 863.50 Matches: 204
Percent Similarity: 55.3% Conservative: 90
Best Local Similarity: 38.3% Mismatches: 187
Query Match: 30.3% Indels: 51
DB: 2 Gaps: 12

```

```

US-10-665-990A-13 (1-1561) x Q7MMB3_VIBVY (1-554)
QY 29 ATATCCCGCATGAAAAACA-----CGAGCCTCATTTCCCTTTTATGCTCTTCTC 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
47 LeuGlyProLeuArgThrPheIleHisLysAlaLeuThrTyrCysSerValMetLeuLeu 66
QY 80 TGTTCATGTTCTTTCATGG-----TTGCCCCCTCGAAGAACCGACGCGAA----- 124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 CysAlaCysThrSerPheGluAsnHisSerProPheAspLysGlnProSerTyrGlnLeu 86
QY 125 -----AGCGTCATTTCAATACTTCCAAACCTGTCTCTCTCTG 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87 GlyTyrGlnAlaAspSerArgLeuSerAlaTyrLeuAsn----- 99
QY 161 GACAAACATCTGCAAAATCCGGCACACCCCTCATACACCGGCTATCCGACATCTACCTG 220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 -----HisHisProGlnAspArgGluAsnLeuThrAlaPhePhePro 113
QY 221 CTCGACGACCCCGACGAAGCCCTTGGCCGCGCGCTTATCGAATCTGCCGAACAC 280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 LeuAspLysGlyHisAspAlaLeuAlaArgLeuAlaLeuIleGluAlaAlaAspLys 133
QY 281 AGCTCGATTGCAATACTACTTTGGCGCAACGACATTTCCGCGAGCTGCTGTTCAC 340
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 ThrLeuAspLeuGlnTyrTyrIlePheArgAsnAspGluThrGlyGlnLeuLeuThrTrp 153
QY 341 CTCATGTACCTTCCCGCAGAACCGCGCTGCGGTAGCCCTGCTGTGGAGCAGACAAAC 400
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
154 ArgLeuPheGluAlaAlaGlnArgGlyValArgValArgIleLeuLeuAspAspMetGln 173
QY 401 ACGCGCGGTGGACGATCTCTGCTCGCTCGACAGCCATCCCAATATCGAAGTGGCG 460
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
174 LysArgAsn---AspGluGlyLeuAlaArgLeuSerAlaHisProAsnIleGlnIleArg 192
QY 461 CTGTTCAACCCCTTCTGCTACGCAATGGCGGCATCTCGCTACCTGACCGACTCCCC 520
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 LeuPheAsnProHisGlnTyrArgThrAlaArgThrLeuAlaMetAlaSerAspPheSer 212
QY 521 CGCTCAACCGCGCATGCAACAACTCTTTACCGCGCCGACACACCGCGCCACCATCTC 580
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 ArgLeuAsnArgArgMetHisAsnLysSerLeuThrAlaAspSerValAlaIleVal 232
QY 581 GGGCGGCGCAATATCGCGCAGCAATCTTCAAAGTGGTGAGGACACCGTTTCCGCCAC 640
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 GlyGlyArgAsnIleGlyAsnGluTyrPheSerValAsnSerProValGluPheGlyAsp 252
QY 641 CTGACATCTCCGACCGCGCAGCGTGTGCGGGAAGTATCGCAGCACTTCGACCGTAC 700
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 PheAspLeuMetLeuTyrGlyAsnSerValGluGlnThrAlaGluGlnPheAspLeuTyr 272
QY 701 TGGGCAAGCCATTTCCGCC-----CACACGGCCACCGGCATCATC 739
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 TrpAsnSerLeuHisAlaValProIleGluTrpLeuThrAspAsnProIleProValThr 292
QY 740 CGCAGCGGCAACATCGGCAAGGCTTTCAGAGCACTCGGATACACACGCAAAATCC--- 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 Glu-----GluGluLeuGlnAla-----TrpLeuLysGluThrGlnLeu 305
QY 797 AGACACGCGCTCTCGCGCTACCGCGAAACCGTCGAACAGTCGCCCTCTTACCAAAAATA 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
306 GluGlnLysPheThrGlnGlyArgTyrAspPheThrGlnLeuGlnLeuTyrGlnGlnPhe 325
QY 857 CAGACGGGCGCATCGACTGGCAGAGCGTCCAAACCGCCTGATCAGCGACACCCCTGCA 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
326 ThrAspLysSerLeuValTrpTyrTrpGlyLysGlyGlnValTrpTyrAspLeuProAsp 345
QY 917 AAAGGAGCTCGACCGCGACCGCGCAAAACCGCCATTGCCGGAGGCTCGAAGACGCGCTC 976
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 Lys-----ValAspThrGlnGluGlnGlnLeuAlaAspAsnLeuThrSerLeuLeu 362
QY 977 AAACAGCCCGGAAAAAGCGTCTATCTGTGTTTACCTTATTCGTCCTACAAATCCGGC 1036
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 ArgThrValLysAspSerLeuValLeuIleSerProTyrPheValProThrGluArgGly 382

```



```

RESULT 24
Q4FTL4_9GAMM
ID Q4FTL4_9GAMM PRELIMINARY; PRT; 570 AA.
AC Q4FTL4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Probable phospholipase D.
GN ORFNames=Psyc_0791;
OS Psychrobacter_arcticum 273-4.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Psychrobacter.
NCBI_TaxID=259536;
[1]
RN NUCLEOTIDE SEQUENCE.
RP
RC STRAIN=273-4;
RG NASA Astrobiology Institute;
RA Ayala-del-Rio H.L., Chain P., Ponder M.A., Di Bartolo G., Ivanova N.,
R Bergolzh P.W., Hauser L., Land M., Bakermans C., Rodrigues D.,
RA Klappenbach J.A., Zarka D., Larimer F., Richardson P., Thomasow M.F.,
RA Tiedje J.M.;
RT "Complete sequence of Psychrobacter arcticum 273-4.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=273-4;
RG
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Iserani S., Chain P., Di Bartolo G., Ivanova N., Hauser L.,
RA Land M., Larimer F., Pitluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000082; AA018644.1; -; Genomic DNA.
SQ SEQUENCE 570 AA; 63466 MW; 0ECC7DAA9ECA3BC1 CRC64;

Alignment Scores:
Pred. No.: 2,27e-39 Length: 570
Score: 852.00 Matches: 196
Percent Similarity: 54.4% Conservative: 99
Best Local Similarity: 36.2% Mismatches: 208

```

```
QY 970 CCGCTCAACAGCGCGAAGAAAGCGTCTATCTGCTTCCACCTATTTCGCTCCCTACAAA 1029
Db 369 rLeuLeuGlySerProSerLysLeuThrIleSerSerTyrPheValProThrLy 389
1030 ATCCGGCCACAGCAGCAGTGGCAAAAGCTGTGTGAGGAGCAGATAGACGTTACCGTCTCGAC 1089
Db 389 sAspGlyValAsnThrLeuAsnAlaLeuAlaGluSerGlyIleGluIleLysIleLeuTh 409
1090 CAATCTGCTACAGGCGAGCGAGTGGCCGCTCATTCGCGCTACGTCGAATAATACCGAAA 1149
Db 409 rAsnSerPheAspAlaThrAspValThrAlaValHisSerGlyTyrSerGlnTrpArgPr 429
1150 ACCGCTGCTCAAAGCGCGCATCAACTCTAGAGCTGCAACCAACCATCGGTCCTCCGCG 1209
Db 429 oSerLeuLeuArgAlaGlyValIleTyrGluLeuLysSerThrAlaSerGluGluLy 449
1210 CACAAAGACAAAGCGCTGACCGGG-----AGCTCCGCTAACCGAGCTGCGATGCCAAAC 1263
Db 449 sArgGluAsnLysLeuTrpLysGlyArgSerGlnSerSerThrSerLeuHisAlaLysAl 469
1264 TTTCATTTGTGAGCGCAACAGCATCTTCATCGGCTCATTCACCTCGACCGCCCGTTCGCG 1323
Db 469 aPheAlaValAspAspTyrGlnValPheIleGlySerTyrAsnValAspProArgSerAl 489
1324 ACGGCTCACTACCGAAATGGGCGTCTCATCGAAAGCGCCCAAAATCGCAGACAGATGGA 1383
Db 489 aAsnIleAsnThrGluMetGlyValIleIleAsnAspAspGluLeuAlaThrGlnLeuHi 509
1384 GCGCACCTCGCGCGATACACACCGCAATACCGCTACCGCGTTACCTCGACAAACACAA 1443
Db 509 sAspAlaLeuSerAspAspLeuLeuAsnGlnAlaTyrGluValLysLeuLeuGluAsnGl 529
1444 CGGCTGCAATGGCAC-----GATCCCGCCACCGCAAAACCTACCGAAGAACCGGA 1497
Db 529 yAsnLeuGlnTrpHisThrMetGluLysAspLysLysValIleTyrAspSerGluProAr 549
1498 AGCCAACTTTGGAACGCATCGCGCAAAATCTATCCCTGCGCCCATCGAAGTTT 1557
Db 549 gValAlaValSerAspHisValTrpLeuThrIleMetSerTrpLeuProIleAspTrpLe 569
1558 ATTA 1561
Db 569 uLeu 570

RESULT 25
QBIM4_XANAC
ID QBIM4_XANAC PRELIMINARY; PRT; 520 AA.
AC QBIM4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Fornighieri E.F., Franco M.C., Greggio C.C., Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.W., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
```

```
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE011929; AAM37716.1; -: Genomic DNA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 57436 MW; 61A2DE5B0349A2FA CRC64;

Alignment Scores:
Pred. No.: 2.73e-39 Length: 520
Score: 850.50 Matches: 206
Percent Similarity: 52.7% Conservative: 80
Best Local Similarity: 37.9% Mismatches: 182
Query Match: 29.8% Indels: 75
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x QBIM4_XANAC (1-520)
QY 56 ATTTCCCTTTATGCTCTCTCTCTGTTTCATGT----- 88
Db 12 IleAlaLeuValLeuValAlaSerAlaLeuSerLeuTyrGlyTyrGlyArgPhe 31
QY 89 -----TCTTCATGGTTGCCCTGAGAGAACGGACGAA 124
Db 32 AlaAspArgGlnArgGlyProValSerHisValLeuProAla----- 45
QY 125 AGCGGTCAATTCAATACCTTCCAAACCTGCTCTCTGGACAAACATC-----CTGCAA 175
Db 46 -----ThrAlaValAlaThrProIleAspLysValAlaProLeuGln 60
QY 176 ATCCGGCACACCCCTCATACACAGGGCTATCCGACATCTACCTGCTCGAGACCCCCAC 235
Db 61 GlnAlaHisAlaAspGln-----ThrGlyMetValIleLeuProAspAsnVal 76
QY 236 GAAGCCCTTGGCGCGCGCCCTTATCGAATCTGCGCAACACAGAGCTCGATTTCGAA 295
Db 77 AspAlaPheAlaValArgAlaLeuThrAlaArgAlaAlaGlyArgSerLeuAspLeuGln 96
QY 296 TACTACATTTGGCGCACACGACATTTCCGGCAGGCTGTGTTCAACCTCATGTACCTGCC 355
Db 97 TyrTyrIleTrpHisAlaAspPheThrGlyAsnLeuLeuHisAsnGluLeuArgAla 116
QY 356 GCAGAACCGCGGTGGGTAGCGCTGCTGGAGCAGACACACACGCGCGGTGGAC 415
Db 117 AlaAspArgGlyValArgValArgLeuLeuLeuAspMetAsnIleHisGlySerAsp 136
QY 416 GATCTCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTC 475
Db 137 SerValLeuAlaAlaLeuAspSerHisProLeuIleGluIleArgLeuPheAsnProThr 156
QY 476 GTCCTACGCAAA-----TGGCGCGCACTCGGCTACCTGACCGACTTCCCGCGCTC 526
Db 157 ArgAlaArgGluGlyThrLeuMetArgGlyValGluMetValLeuArgMetPheSerIle 176
QY 527 AACCGCCGATGCACAAACAAATCTTTACCGCGCAACACCGCGCCACCATCTACTCGCGGA 586
Db 177 AsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgIleAlaValValGlyGly 196
QY 587 CGCAATATCGCGCAGCAATATCTTCAAGTCGTGAGGACACCGTTTTTCGCCACCTGGAC 646
Db 197 ArgAsnValGlyAspGluTyrPheAspAlaAlaArgAspThrAsnPheMetAspMetAsp 216
QY 647 ATCTCCGCCACCGCGAGCGTGTGCGGAGAGTATCGACGACTTCGACCGCTACTCTGGGA 706
Db 217 AlaAlaLeuMetGlyProAlaValGlyGluGlnValPheAspAlaTyrTrpAsn 236
```

Qy	707	AGCCATTCGGCCGACAAACGCCGCGCATCATCCGCAGC-----	745
Db	237	SerProAsnAlaLeuProLeuSerAlaLeuValThrAlaLysProGlnAlaLeuGluAla	256
Qy	746	-----GCCAATCGGCAACGGGCTTCAAGCACTCGGATCAACGACGAACATCCAGA	799
Db	257	LeuArgGlySerLeuAspAlaGlyMetAlaSerValArgAlaHisPro-----	272
Qy	800	CACGCGCTCTCGCTACCGCGAAGACCGCTCGAAACAGTCGCGCCCTCTACCAAAAATACAG	859
Db	273	-----TyrValGluArgLeuLysArgSerProSerValArgGluLeuMet	287
Qy	860	ACGGGA-----CGCATCGACTGGCAGAGCTCCAAACCCCGCTGATCGACGACACCCCT	913
Db	288	GlnGlyAspArgProValHisTrpValSer--GlnAlaArgIleValSerAspProPro	306
Qy	914	GCAAAAGGA-----CTCGACCGCAGCCGCCCAACCGCCGATGTCGCGGAGG	961
Db	307	GluLysAlaGluGlyAlaProProGlnAlaAspTrpMetThrProValLeuIleGlyGlu	326
Qy	962	CTGCAAGACGGCTCAACACAGCCCGAAAAAGCGTCTATCGTTTTCACCCCTATTTCGTC	1021
Db	327	MetAlaHisAla-----GlnArgGluLeuLysValIleSerProTyrPheVal	342
Qy	1022	CCTACAAAATCCGGCAGACGCACTGGCAAAATCGTGTGAGGACGGCATAGAGCTTACC	1081
Db	343	ProGlyAspGluGlyMetArgTrpIleGlyHisLeuArgGlnArgAsnValArgValSer	362
Qy	1082	GTCCTGACCAACTCGCTACAGGACCGACGCTTCGCGCGTCCATTCGGGTACGTCAAA	1141
Db	363	IleLeuThrAsnSerLeuAlaAlaAsnAspValValAlaValHisSerGlyTyrAlaGly	382
Qy	1142	TACCGAAAACCGCTCTCAAGCCGGCATCAAACTCTACGAGCTGCACCCCAACCATGCC	1201
Db	383	TyrArgValProLeuLeuGlnLeuGlyValArgLeuHisGluLeuLysPro-----	399
Qy	1202	GTCCCCGCCACAAAAGCAAGGCTGCACGGCAGCTCCGTAACACGAGCTGCATGCCAAA	1261
Db	400	---MetGlyLysProAspGlySerLeuPheGlySerSerGlyAlaSerLeuHisThrLys	418
Qy	1262	ACCTTCATTGTGGACGGCAACGCGATCTTCATCGGCTCATTCACACCTCGACCCCGTTCC	1321
Db	419	AlaPheValValAspAspSerSerGlyPheIleGlySerPheAsnLeuAspProArgSer	438
Qy	1322	GCACGGCTCAATACCGAAATGGGCGTCGTATCGAAAGCCGCCAAAATCGCAGAACAGATG	1381
Db	439	MetAsnLeuAsnThrGluMetGlyLeuLeuPheAsnAspArgValValThrAlaGluLeu	458
Qy	1382	GAGCGCACCCCTC---GCCGATACACACCCGGAATACGCTTACCCGGTTACCTCGACAAA	1438
Db	459	GluArgLeuTyrAsnHisLysValSerAlaProValSerTyrArgValThrLeuAspArg	478
Qy	1439	CACAACCGCTGCAATGCACGATCCCGCACCCCGA-----AAAACCTACCCGACGAA	1492
Db	479	---GlyGluLeuArgTrpHisAspAspAlaAlaGlnProProGluValIleTrpSerArgGlu	497
Qy	1493	CCCGAACCCAAACTTTGGAAACGCGATCCGCGCAAAAAATCCCTATCCCTGCTGCCCATCGAA	1552
Db	498	ProAlaAlaSerValTrpArgArgGlyAlaAlaValMetGlyTyrTrpLeuProLeuGlu	517
Qy	1553	GGTTTATTA	1561
Db	518	SerGlnLeu	520

RESULT, T 26

RESULI 26
06MOL7 BDFRA

Q6MQL7_BDEBA
ID Q6MOL7 E

ID	Q6MQL7_E
AC	Q6MOL7:

AC DT 05-JUL-2005

DT 05-JUL-2005

DT 05-JUL-20

DE Putative

4
5
6
7
8
9
10
11
12

GN	OrderedLocusNames=Bd0448;
OS	Bdellovibrio bacteriovorus.
OC	Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OX	Bdellovibrionaceae; Bdellovibrio.
NCBI_taxID=959;	
EN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX	PubMed=14752164; DOI=10.1126/science.1093027;
RA	Rendulich S., Jagtap P., Rosinus A., Eppinger M., Baar C.,
RA	Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
PA	Sockett R.E., Schuster S.C.;
RT	"A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a genomic perspective.";
RL	Science 303:689-692(2004).
DR	EMBL; BX842647; CAE78430.1; -; Genomic DNA.
DR	GO; GO:0003824; F:catalytic activity; IEA.
DR	GO; GO:0008152; P:metabolism; IEA.
DR	InterPro; IPR001736; PLD.
DR	Pfam; PF00614; PLDC; 2.
DR	SMART; SM00155; PLDC; 2.
DR	PROSITE; PS50035; PLD; 2.
KW	Complete proteome; Signal.
FT	SIGNAL 1 20 Potential.
SQ	SEQUENCE 530 AA; 59501 MW; 4C810BFADF934F6F CRC64;
Alignment Scores:	
Pred. No.:	3.11e-39 Length: 530
Score:	849.50 Matches: 203
Percent Similarity:	55.5% Conservative: 84
Best Local Similarity:	39.3% Mismatches: 205
Query Match:	29.8% Indels: 25
DB:	2 Gaps: 12
US-10-665-990A-13 (1-1561) x Q6MQL7_BDEBA (1-530)	
Qy	56 ATTTCCTTTATAGCGTCTCTGTTGCATGTCTTCATGGTTGCCCTCGGAAGA 115
Db	24 ILeProIleLeuLeuGlyLeuSerAlaglyCysAlaSer---LeuProAlaAsnThrAsn 42
Qy	116 CGG-----ACGGAAACGGTCATTCAATCTTCCAACCTGCTCCTCGACAACATC 169
Db	43 ArgProTyrrSerGluAlaLeuProProAspProlThyThrProLeuAlaGlnAlaThrGlu 62
Qy	170 CTGCAAAATCGGGCACACCCCTCATAAACACGGGTATCCGCATCTACCTGCTCGACGAC 229
Db	63 ThrGlnIleLysGlyHisPro-----GlyGlnSerGlyPhenylrProLeuAlaSer 79
Qy	230 CCCACGAAGCCCTTGCGCGCGCGCGCCCTTATCGAATCTGCCGAACACAGCCTCGAT 289
Db	80 GlyLeuGluAlaLeuValAlaArgMetAlaAlaValValSerAspArgSerileAsp 99
Qy	290 TTGCAATACTACATTTTGGCGCAACAGCATTTCCGCGAGGTGCTGTTCAACCTCATGTAC 349
Db	100 LeuGlnTyrrTrileTrpGluAsnAspLeuthrGlyArgMetLeuHisGluValLeu 119
Qy	350 CTTGCCGCAGAACCGCGGTGCGGTAGCTGCTGTTGGAGCAGCAAACAACGCGCGG 409
Db	120 ArgAlaAlaaspargGlyValargvalargValLeuLeuAspAspLeuAsnGlnSerArg 139
Qy	410 TTGACAGATCTCCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCAAC 469
Db	140 TyrGlulysGlyLeuAlaIleLeuAspyrHisProAsnValGluValArgMetAlaAsn 159
Qy	470 CCCTTGCTCTACGCAATATGGCGGCACTCGGCTACTGTGACCGACTTCCCCCGCTCAAC 529
Db	160 ProPheAlaGlyArgThrTrpLysIleLeuAspAlaMetArg---PheSerThrValAsn 178
Qy	530 CGCGCGATGCACAACAAATCTTTACCGCGACACCAGCGCCACCATACTCGGCGGACGC 589
Db	179 ArgGmethHisasnlysvalPheValValAspAsnGlnThrAlaIleValIglyGlyArg 198
Qy	590 AATATCGGCGACGAATACTTCAAAGTCGGTAGGACACCGTTTTTCGCCGACCTGGACATC 649

```
Db 199 AsnIleGlyAspGluTyrPheGlyAlaSerGluGluMetAsnPheGlyAspPheAspLeu 218
Qy 650 CTCGCCACCGCAGCGTCGTGGGAAAGTATCGCACGACTTCGACCGCTACTGTGGCAAGC 709
Db 219 TrpAlaIleGlyProValValGlnGluLeuSerArgHisPheAspThrTyrTrpAsnSer 238
Qy 710 CATTCCGCCCAACAGCCACCGGCATCATCGCAGCGGCAACATCGGCAAGGGTCTTCAA 769
Db 239 GluIleAlaTyrProIleSerValLeuValLys-----GlyPheLysProThrLeuGlu 256
Qy 770 GCATCTCGGATACACAGCAGAAACATCCAGACACGCGCTC-----CTGCGC 814
Db 257 AspLeuGlnLysLeuLysAspAlaAlaAlaIleThrGluAlaGluLysThrGln 276
Qy 815 TACCGCGAAACCGTCGACAGCTCGCCCTCTACCAAAAATA-----CAGACGGGAGCG 868
Db 277 TyrAlaAspAlaLeuLysGluThrProIleValLysLysPheThrHisGluProMetLys 296
Qy 869 ATCGACTGGCAGAGGTCCTCAACCGCTGTATCAGCGACACCCCTGCAAAA---GGACTC 925
Db 297 LeuTyrTrpGlyLysAlaAsp-----ValValMetAspProProGluLysPheHisGln 314
Qy 926 GACCGCGACCGCGCAACCGCGATTCGCGGGAGGCTGCAAGAGCGCTCAACACAGGCC 985
Db 315 AspSerLysAspGlnValAspAsnLeuAlaArgGlnLeuTyrProLeuIleGluLysThr 334
Qy 986 GAAAAAGCGCTATCTGTTTTCACCCCTATTTCCTCTACAAAATCCGGCACACAGCA 1045
Db 335 GluLysGluLeuIleLeuValSerProTyrPheValProGlyLysLysGlyValLysPhe 354
Qy 1046 CTGGCAAACTGGTGCCAGGAGCGATAGACGTTACCTGCTGACCACTCGTACAGCGC 1105
Db 355 PheLysHisLeuAsnAspArgGlyValGlnSerLeuValLeuThrAsnSerLeuAlaSer 374
Qy 1106 ACCGAGCTGCGCGCTCATTCGCGGTACGTCAAAATACCGAAACCGCTGCTCAAAGCC 1165
Db 375 SerAspValAlaThrValPheSerGlyTyrLysGlyTyrArgLysAspLeuLeuAspGln 394
Qy 1166 GGCATCAAACTCTACGAGCTGCAACCCCAACCATCGCTGCCCGCCGACAAAAGACAAAGGC 1225
Db 395 GlyValSerLeuTyrGluLeuLysProAsn---SerProThrThrThrProLysLysAsn 413
Qy 1226 CTGACCGCGCAGCTCGTTAAACAGC-----CTGCATGCCAAACCTTCATTGTGGACGGC 1279
Db 414 ArgValGlySerSerPheSerAlaGlyLeuHisGlyLysIlePheValPheAspArg 433
Qy 1280 AAACGCATCTTCATCGCGTCTATTCAACCTGACCCCGCTTCGCGACGGCTCAATACCGAA 1339
Db 434 LysLysValPheValGlySerMetAsnLeuAspProArgSerAlaThrLeuAsnSerGlu 453
Qy 1340 ATGGCGCTGTCATCGAAGCCCCAAAATCGCAGAACATGAGCGCACCCCTCGCGCAT 1399
Db 454 MetGlyValValValAspSerProGluLeuAlaGluMetIleSerThrAsnLeuIleAla 473
Qy 1400 ACCACACCGAATACGCGCTACCGGTTTACCTCGACAAACACACACCGCTGCAATGG--- 1456
Db 474 HisLeuArgArgAspSerTyrGlnValLeuLeuAsnGluLysAsnAsnLeuIleTrpLys 493
Qy 1457 -----CAGCATCCCGCCACCGAAAAACCTTACCGGAACCGAACCGAACCAAACTTTGG 1510
Db 494 ThrThrAspAspCysGlyLeuGluHisValPheSerLysAspProGluThrSerTrpTrp 513
Qy 1511 AAACGATCGCGCAAAAATCTATCTCTGTCGCCCATCGAAGGTTTATTA 1561
Db 514 LysArgPheLysAlaSerLeuSerAlaIlePheIleProGluSerTrpLeu 530

RESULT 27
Q5QUY8_IDILO
ID Q5QUY8_IDILO PRELIMINARY; PRT; 464 AA.
AC Q5QUY8;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
```

```
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=lu1444;
OS Idiomarina loihiensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Idiomarinaceae; Idiomarina.
OX NCBI_TaxID=135577;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=L2-TR / DSM 15497 / ATCC BAA-735;
RX PubMed=15596722; DOI=10.1073/pnas.0407638102;
RA Hou S., Saw J.H., Lee K.S., Freitas T.A., Belisle C., Kawarabayasi Y.,
RA Donachie S.P., Pikina A., Galperin M.Y., Koonin E.V., Makarova K.S.,
RA Omelchenko M.V., Sorokin A., Wolf Y.I., Li Q.X., Keum Y.S.,
RA Campbell S., Denery J., Aizawa S.-I., Shibata S., Malahoff A.,
RA Alam M.;
RA "Genome sequence of the deep-sea gamma-proteobacterium Idiomarina
RT loihiensis reveals amino acid fermentation as a source of carbon and
RT energy.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:18036-18041(2004).
DR EMBL; AE017340; AAV82284.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 464 AA; 53557 MW; 59A8BFA13DEB45F7 CRC64;

Alignment Scores:
Pred. No.: 4,81e-39 Length: 464
Score: 846.00 Matches: 186
Percent Similarity: 56.6% Conservative: 76
Best Local Similarity: 40.2% Mismatches: 173
Query Match: 29.7% Indels: 28
DB: 6

US-10-665-990A-13 (1-1561) x Q5QUY8_IDILO (1-464)
Qy 218 CTGCTCGACGCCCCACGAGAGCTTCCTGCCCGCCGCGCTTATCGAATCTGCCGA 277
Db 15 LeuLeuGluAsnAlaHisArgAlaLeuIleGluArgValArgLeuLeuArgGlnAlaArg 34
Qy 278 CACAGCTCGATTGTCGAATACTACTATTGGCGCAACGACATTCGGCAGGCTGTCTTC 337
Db 35 HisHisIleAlaLeuGlnTyrTyrLeuTrpArgProAspSerSerGlyLeuThrLeuLeu 54
Qy 338 AACCTCATGTACCTTCGCGAGAACGCGCGCTGCGGTACGCTGCTGTGGACGACAAAC 397
Db 55 LysGluLeuLeuAspAlaValGluArgGlyValArgValAspLeuLeuLeuAspLeu 74
Qy 398 AACACGCGCGGGTTGGACGATCTCTGCTGCCCTCGACAGCCATCCCAATATCGAAGTG 457
Db 75 HisSerLysProLeuGluProLeuLeuArgAspLeuSerSerArgAlaAsnPheAsnIle 94
Qy 458 CGCTGTGTCAACCCCTTCCTCCTACGCAATTCGGCGGCACTCGGCTACTGACCGACTTC 517
Db 95 ArgLeuPheAsnProPheLysHisArgArgTrpArgTrpLeuAsnTrpLeuThrAspPhe 114
Qy 518 CCCCGCTCAACCGCGCATGCACAAATCTTTTACCGCGCAACACCGCGCCACCATTA 577
Db 115 LysArgMetAsnArgArgMetHisAsnLysSerMetLeuValAspAlaGluValGlyIle 134
Qy 578 CTCGGCGGACGCAATATCGGCGAGCAATCTTCAAAAGTCGGTGAGGACACCGTTTTTCGC 637
Db 135 ValGlyGlyArgAsnValGlyAsnGluTyrPheGlyThrHisAlaGlyGlnLeuPheSer 154
Qy 638 GACCTGGACATCTCTCGCCACCGGAGCTGCTCGCGGAAGTATCGCAGCTTCGACCGC 697
Db 155 AspLeuAspValIleAlaLysGlyLysIleValLysThrLeuLeuAspTrpArgGln 174
Qy 698 TACTGGGCAAGCCATTTCGCCCCACAAACGCCACGCGCATCATCCGCGAGCAACATCGGC 757
```



```
Qy 854 ATACAGACGGGACGC-----ATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCGAC 907
Db 286 LeuMetGlnGlyAsnArgProValHisTrpValGln---GlnAlaArgIleValAlaAsp 304
Qy 908 ACCCTCGCAAAAGCA-----CTCGACCGGACCGCGCAAAACCGCGCTGATTCGCC 955
Db 305 ProProGluGlyAlaGluGlyAlaProProGlnProAspTrpMetThrProIleLeuIle 324
Qy 956 GGGAGGTGCTGAAGACGGCTCAAAACAGCCGCAAAAGCGTCTATCTGTTTACCTCAT 1015
Db 325 GlyLeuMetAlaAsnAla-----ArgArgGluLeuLeuLeuLeuLeuSerProTyr 340
Qy 1016 TTCTGCTCTCAAAATCCGGCACAGACGACTGCGCAAACTGGTGCAGGCGGCATAGAC 1075
Db 341 PheValProGlyAspAspGlyMetArgTrpIleGlyLeuArgGlyAspValArg 360
Qy 1076 GTTACCGTCTGACCAACTCGCTCAGCGGACCGACGTTGGCGCGTCCATTCCTCGGGTAC 1135
Db 361 ValSerValLeuThrAsnSerLeuAlaAlaAsnAspValValAlaValHisSerGlyTyr 380
Qy 1136 GTCAAAATACCGAAACCGCTCTCAAAACCGCGCATCAAACTCTACGAGCTGCAACCCAAC 1195
Db 381 AlaAspTyrArgIleProLeuLeuGlnGlyValArgLeuHisGluLeuLeuPro--- 399
Qy 1196 CATCCGCTCCCGCCCAAAAGACAAAGGCTGACCGGACGCTCGGTAAACGAGCTGCAT 1255
Db 400 -----MetGlyArgProAspGlySerLeuPheGlySerSerGlyAlaSerLeuHis 416
Qy 1256 GCCAAAACCTTCATTGTGGACGGCAAAACGCGATCTTCATCGGCTCATTCACCTCGACCCC 1315
Db 417 ThrLysAlaPheValValAspSerAlaGlyPheIleGlySerPheAsnLeuAspPro 436
Qy 1316 CGTTCGCGACGGCTCAATACCGAAATGGCGTCTCATCGAAAGCCCAAAATCCGAGAA 1375
Db 437 ArgAlaMetAsnLeuAsnThrGluMetGlyLeuLeuPheGluAspArgThrValThrAla 456
Qy 1376 CAGATGAGCGCACCCCTC---GCCGATACACACCCCGAATACGCTACCGGTACCTC 1432
Db 457 GluLeuGluArgLeuTyrAsnHisLysValSerAlaProValSerTyrArgValThrLeu 476
Qy 1433 GACAAACACACACCGCTGCAATGCGATGCGACGATCCCGCCACCCGA-----AAACCTACCCG 1486
Db 477 GluGln---GlyGlyLeuArgTrpHisAspAlaAlaGlnProProAlaValTrpThr 495
Qy 1487 AACGAACCGGACCAACTTTGGAAACGATCGCGCAAAATCTATCCCTGCTGCC 1546
Db 496 ArgGluProAlaAlaSerValTrpArgGlyAlaAlaLysValLeuGlyTrpLeuPro 515
Qy 1547 ATCGAAGGTTTATTA 1561
Db 516 LeuAspSerGlnLeu 520

RESULT 29
Q8P7A5_XANCP
ID Q8P7A5_XANCP PRELIMINARY; PRT; 520 AA.
AC Q8P7A5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1];
RP NUCLEOTIDE SEQUENCE
RC STRAIN=ATCC 33913 / NCPB 528;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
```

```
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE012383; AAM41978.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 57165 MW; 8849CFED240410A3 CRC64;

Alignment Scores:
Pred. No.: 7.16e-39 Length: 520
Score: 843.00 Matches: 196
Percent Similarity: 55.1% Conservative: 71
Best Local Similarity: 40.4% Mismatches: 174
Query Match: 29.6% Indels: 44
DB: 2 Gaps: 12

US-10-665-990A-13 (1-1561) x Q8P7A5_XANCP (1-520)
Qy 170 CTGCAAAATCCGGCACACCCCTCATACACAGCGGTATCCGACATCTACCTGCTCGACGAC 229
Db 59 LeuGlnGlnAlaHisAla-----GlyLysThrGlyMetValIleLeuSerAsp 74
Qy 230 CCCCGAAGCCCTTGGCGCGCGCGCTTATCGAATCTGCGGACACACAGCTCGAT 289
Db 75 AsnIleAspAlaPheAlaValArgAlaLeuThrAlaArgAlaAlaGlyArgSerLeuAsp 94
Qy 290 TTGCAATACTACATTTTCGCGCAACGACATTTCCGGCAGGCTGTCTCAACTCATGTAC 349
Db 95 LeuGlnTyrTrpIleThrHisAlaAspPheThrGlyAsnLeuLeuHisAsnGluLeuLeu 114
Qy 350 CTTGCGCGCAGAACGCGCGTACGCGTACGCTGTGTGGACGACAAACACACGCGCGGG 409
Db 115 ArgAlaAlaAspArgGlyValArgValArgLeuLeuLeuAspMetAsnIleHisGly 134
Qy 410 TTGGACGATCTCTGCTCGCGCTCGACAGCATCCCAATATCGAAGTGGCGCTGTTCAAC 469
Db 135 SerAspSerValLeuAlaAlaLeuAspSerHisProLeuIleGluIleArgLeuPheAsn 154
Qy 470 CCCTGCTCTCTACGCAAA-----TGGCGCGCACTCGGCTACCTGACCGACTTCGCC 520
Db 155 ProThrArgAlaArgGluGlyThrLeuMetArgGlyValGluMetValLeuArgMetPhe 174
Qy 521 CGCTCAACCGCGCATGCGACACAAATCTTTACCGCGCGGACCAACCGCGCCCATCTC 580
Db 175 SerIleAsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgIleAlaValVal 194
Qy 581 GCGCGGACGCAATATCGCGGACGAATATCTTCAAGTCGGTGGAGGACACCGTTTTCGCCG 640
Db 195 GlyGlyArgAsnValGlyAspGluTyrPheAspAlaAlaArgAspThrAsnPheMetAsp 214
Qy 641 CTGGACATCTCTCGCCACCGCGCGTGTGTGGCGAAGTATCGCACGACTTCGCGGTAC 700
Db 215 MetAspAlaAlaLeuIleGlyProThrValArgGlnSerGluGlnValPheAspAlaTyr 234
Qy 701 TGGCAAGCCATTCGCCGCCCAACACCGCCGCGCATCATCCGACG-----745
Db 235 TrpAsnSerProAsnAlaLeuProLeuAlaAlaLeuValThrAlaLysProGlnAlaLeu 254
```

```
QY 746 -----GGCAACATCGGCAAGGTTCTTCAAGACATCGGATACAAACGACGAAACA 793
Db 255 GluGlnLeuArgGlySerLeuAspAlaGlyMetAlaSerValGlyAlaHisPro----- 272
QY 794 TCCAGACACGGGCTCTCGGCTACGGGAAACCGTCGAAACAGTCCGCCCTCTACCAAAA 853
Db 273 -----TyrValGluArgLeuArgGlnSerProSerValHisAla 285
QY 854 ATACAGACGGGACGC-----ATCGACTGGCAGAGCGTCCAAACCCCGCTCATCAGCAGC 907
Db 286 LeuMetGlnGlyAsnArgProValHisTrpValGln---GlnAlaArgIleValAlaAsp 304
QY 908 ACCCTGCAAAAGGA-----CTCGACCGCAGCAGCCGCCCAACCGCCCATGCC 955
Db 305 ProProGluLysAlaGluGlyAlaProProGlnProAspTrpMetThrProIleLeuIle 324
QY 956 GGGAGGCTGCMAGCGGCTCAACACGCCCGAAAGCGTCTATCTGGTTTACCCCTAT 1015
Db 325 GlyGluMetAlaAsnAla-----ArgArgGluLeuLysLeuIleSerProTyr 340
QY 1016 TTCCTCCTACAAATCCGGCAGACGCACTGCGCAAACTGGTGCAGGACGGATAGAC 1075
Db 341 PheValProGlyAspAspGlyMetArgTrpIleGlyLeuArgLysArgAspValArg 360
QY 1076 GTTACCGTCTGCACTCGCTACAGCGACCGCGTTCGCGCGTCCATTCGGGTAC 1135
Db 361 ValSerValLeuThrAsnSerLeuAlaAlaAsnAspValValAlaValHisSerGlyTyr 380
QY 1136 GTCMAATACGAAACCGCTCTCAAGCGCGGATCAAACTCTACGAGTGCACCCCAAC 1195
Db 381 AlaAspTyrArgIleProLeuLeuGlnGlnGlyValArgLeuHisGluLeuLysPro--- 399
QY 1196 CATGCCGTCCCGCCACAAAGACAAAGGCTGACCGGAGCTCCGTAACACCGCTCAT 1255
Db 400 -----MetGlyArgProAspGlySerLeuPheGlySerGlyAlaSerLeuHis 416
QY 1256 GCCAAACCTTCATTGTGGACGGCAACACGCATCTTCATCGGCTCATTCACACCTCGACCCC 1315
Db 417 ThrLysAlaPheValValAspAspSerAlaGlyPheIleGlySerPheAsnLeuAspPro 436
QY 1316 CGTTCCGACGGCTCAATACGAAATGGCGTGTCTATCGAAAGCCCAAAATCGCGAA 1375
Db 437 ArgAlaMetAsnLeuAsnThrGluMetGlyLeuLeuPheGluAspArgThrValThra 456
QY 1376 CAGATGAGCGCACCCCTC---GCCGATACACACCCGCAATACGCTACCGCTTACCCTC 1432
Db 457 GluLeuGluArgLeuTyrAsnHisLysValSerAlaProValSerTyrArgValThrLeu 476
QY 1433 GACAAACACACCGCTGCAATGGCAGATCCCGCCACCCGA-----AAACCTACCCG 1486
Db 477 GluGln---GlyGlyLeuArgTyrHisAspAlaAlaGlnProProAlaValTrpThr 495
QY 1487 AACGAACCCGAAGCCAACTTTGAAACGATCGCGCAAAATCTCTATCCCTGCTGCC 1546
Db 496 ArgGluProAlaAlaSerValTrpArgGlyAlaAlaLysValLeuGlyTrpLeuPro 515
QY 1547 ATCGAAGGTTTATTA 1561
Db 516 LeuAspSerGlnLeu 520

RESULT 30
Q6G185_BARQU
ID Q6G185 BARQU PRELIMINARY; PRT; 523 AA.
AC Q6G185
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=BQ01680;
OS Bartonella quintana (Rochalimaea quintana).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
```

```
OX NCBI_TaxID=803;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=roulouse;
RX PubMed=15210978; DOI=10.1073/pnas.0305659101;
RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvel M.,
RA La Scola B., Holmberg M., Andersson S.G.E.;
RT "The louse-borne human pathogen Bartonella quintana is a genomic
RT derivative of the zoonotic agent Bartonella henselae.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721 (2004).
DR EMBL; BX897700; CAP25671.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 523 AA; 59265 MW; BDDCE70FC51BDC16 CRC64;

Alignment Scores:
Pred. No.: 4,93e-38 Length: 523
Score: 828.00 Matches: 200
Percent Similarity: 57.0% Conservative: 64
Best Local Similarity: 43.2% Mismatches: 161
Query Match: 29.0% Indels: 38
DB: Gaps: 13

US-10-665-990a-13 (1-1561) x Q6G185_BARQU (1-523)
QY 236 GAAGCCCTTCCGCCCGCGCCCTTATCGAATCTGCCGACACACAGCTCGATTTCGAA 295
Db 78 AsPalapheCysValArgAlaIleGlyAlaAlaGlnAlaGlyArgSerLeuAspLeuWet 97
QY 296 TACTACATTTGGCGCAACGACATTTCCGGCAGCTGTGTTCACCTCATCTACTTTGCC 355
Db 98 Ty-TyrIleTrpAsnAspLeuThrGlyArgLeuLeuSerGluIleValGluAla 117
QY 356 GCAGAACGGCGCTGGCTGACCGCTCTGCTTGGACGACAAACACACCGCGGGTTGGAC 415
Db 118 AlaAspArgGlyValArgValArgLeuLeuAspAspIleAsnAlaAlaArgAsp 137
QY 416 GATCTCTGCTGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCACCCCTTC 475
Db 138 ProAlaTyrIleAlaLeuAspLysHisProArgIleGluValArgMetPheAsnProGly 157
QY 476 GTCCTACGCAATATGG-----CGCGCACTCGGCTACCTCGACCGACTTCCCGCGCTC 526
Db 158 ArgSerArgLysGlyGlyLeuArgArgGlyLeuGluIleIleLeuArgAlaIleThrVal 177
QY 527 AACCGCGCATGCACAAACAAATCTTTACCGCGACAAACCGCGCCACCATCTCGCGGGA 586
Db 178 ThrArgArgMetHisAsnLysAlaPheIleValAspGlyArgValAlaPheValGlyGly 197
QY 587 CGCAATATCGGCAACGAATACTTCAAAGTCGGTGAGGACACCGCTTTTCGCGCACCTGGAC 646
Db 198 ArgAsnLeuAlaAspSerTyrPheAspAlaGlyGluGluSerHisPheArgAspLeuAsp 217
QY 647 ATCCTCGCCACCGCAGCGTGTGGCGAAGTATCGCACACTTCGACCCCTACTGGCCA 706
Db 218 LeuMetLeuIleGlyProSerValLysLysValGluThrIlePheAspAspPheTrpAsn 237
QY 707 AGCCATTCCGCCCAACGCCCGCATCATCCGCGGGGCAACATCGGCAAGGCTCTT 766
Db 238 SerAlaValValLeuProIleHisThrLeuVal-----ValProLysSerThr 253
QY 767 CAAGCACTCGGATACAAACGACGAAACATCCAGACACGCGCTCTCTGGCTACCGCAAAAC 826
Db 254 SerAspLeuGlyTyrTrpLysGluLysLeuArg-----LysPheArgAspSer 269
QY 827 ---GTCGAACAGTCGCGCCCTC---TACCAAAA-----ATA 856
```


Db 309 GluLysValLeu---ArgLysLysAlaGluAsnTrpLeuMetLysAlaLeuSerGlnVal 327
Qy 974 CTCAAACAGCCGAAAGAGCGTCTATCTGGTTTCACCCCTATTTCGTCCTTCAAAATCC 1033
Db 328 IieGluAspAlaGlnLysThrValGlnIleThrSerProTyrPheValProGlyLysVal 347
Qy 1034 GGCACAGACGCACTGGCAAAACTGGTCAGGACGCGATAGACGTTTACCGTCTCGACCAAC 1093
Db 348 GlyThrGlnAsnPheSerAsnLeuValSerLysGlyValAspValLysIleLeuThrAsn 367
Qy 1094 TCGCTACAGGACGCGAGTGGCCCGCTCATTCGGCTACGTCAGTCAATACCGAAACCG 1153
Db 368 SerLeuAlaAlaThrAspValAlaLeuValHisGlyGlyTyrValProTyrArgLysMet 387
Qy 1154 CTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATCGCGTCCCGCCACA 1213
Db 388 LeuLeuLysSerGlyValLysLeuTyrGluLeuLysProGlyGlyAsnMetHisGlyLeu 407
Qy 1214 AAAGACAAAGCCCTGACCGGAGCTCCGTAAACCGCTGCAATGCAAAACCTTCATTGTG 1273
Db 408 Arg-----LeuPheArgSerSerLysAlaSerLeuHisThrLysAlaPheLeuVal 424
Qy 1274 GACGGCAACGCATCTTCATCGGCTCATTAACCTCGACCCCGCTTCGCGACGGCTCAAT 1333
Db 425 AspArgLysThrAlaPheIleGlySerLeuAsnPheAspProArgSerAlaAlaLeuAsn 444
Qy 1334 ACCGAAATGGCGTGGTCATCGAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTC 1393
Db 445 ThrGluMetGlyIleLeuPheGluCysAlaProIleThrThrArgLeuAspLeuPhe 464
Qy 1394 GCC---GATACACACCCGATACCGCTTACCGCTTACCGCTGCAACCAACACCGCGCTG 1450
Db 465 AlaGluGluThrThrGlyGluMetSerTyrHisLeuArgLeuAsnAspAsnArgIle 484
Qy 1451 CAATGGCAGATCCCGCACCGCCGAAACCTTACCGC-----AACGAAACCGGACCAAA 1504
Db 485 TyrTrpAspPheIleGluAsnGluLysGlnTyrSerValAsnTyrGluProGluSerAsn 504
Qy 1505 CTTTGGAAACGATCGCGCAAAATCTATCCCTGCTGCGCCATCGAAGGTTTATTA 1561
Db 505 PheTrpArgAlaPheAlaLysIleIleSerTrpLeuProIleGluSerGlnLeu 523

RESULT 32

Q98NZ3_RHILO
ID Q98NZ3_RHILO PRELIMINARY; PRT; 466 AA.
AC Q98NZ3
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Phospholipase D-family protein.
GN OrderedLocustNames=mlr9675;
OS Rhizobium loti (Mesorhizobium loti).
OG Plasmid pMdB.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003017; BAB54862.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.
DR SMART; SMO0155; PLDC; 2.
DR PROSITE; PS0035; PLD; 2.
KW Complete proteome; Plasmid.
SQ SEQUENCE 466 AA; 51793 MW; 4693C68F20F94F3F CRC64;

Alignment Scores:

Pred. No.: 3,57e-37 Length: 466
Score: 812.50 Matches: 197
Percent Similarity: 55.4% Conservative: 63
Best Local Similarity: 42.0% Mismatches: 176
Query Match: 28.5% Indels: 33
DB: 2 Gaps: 12

US-10-665-990A-13 (1-1561) x Q98NZ3_RHILO (1-466)

Qy 200 GGGGTATTCGACATCTACTGCTCGACGACCCCAAGCCCTTGGCGCCGCGCCGCC 259
Db 16 GlyLysSerGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 35
Qy 260 CTTATCAATCTCGCGAACACAGCCTCGATTGCAATCTACATTTGGCGCAACGACATT 319
Db 36 AlaAlaArgGlyAlaValArgThrLeuAspLeuMetTyrTyrLeuTrpHisAspHis 55
Qy 320 TCGCGCAGGCTGCTGTTTCAACCTCATCTGCTGCGCAGAACCGCGCTGCGCTACGC 379
Db 56 ThrGlyArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 75
Qy 380 CTGCTGTTGGACGACAAACACGCGCGGTTGGACGATCTCTGCTGCGCTCGACAGC 439
Db 76 MetLeuLeuAspValAsnProArgLysSerAspAlaAlaTyrLeuAlaLeuAsnAsn 95
Qy 440 CATCCCAATATCGAAGTGGCCCTTCAACCCCTTCTGCTGCTGCTGCTGCTGCTGCTG 490
Db 96 HisProAsnIleGluLeuArgLeuPheAsnProSerGlyIleArgAlaArgGlyLeuMet 115
Qy 491 CGGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 550
Db 116 ArgGlyAlaGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 135
Qy 551 TTTTACCGCCGACAAACCGCGCCACATCTCGGCGGACGCAATATCGGCGACGAATCTTC 610
Db 136 TrpIleAlaAspSerIleAlaIleValGlyArgAsnValGlyArgAspAlaTyrPhe 155
Qy 611 AAAGTCGGTGAGGACACCGTTTTCGCGACCTCGACATCTCGCCACCGCGAGCTGCTC 670
Db 156 AspAlaAlaGlu---ThrAsnPheArgAspLeuValLeuValLeuGlyProAlaVal 174
Qy 671 GCGCAAGTATCGCACGACTTTCGACCGCTACTGCGCAAGCCATTCCGCCCAACGCGCACG 730
Db 175 GlnGlnThrAlaArgIlePheGlnThrPheTrpAla-----CysGlnAspAlaLys 191
Qy 731 CGCATCATCGCAGCGCAACATCGGCAAGGCTTCAAGCGTCTTCAAGCATCGGATAC-----AAC 784
Db 192 ProIleGlyGluLeuGlyAlaAlaAlaProGlySerHisAlaProAsnPheGluArgSer 211
Qy 785 GACGAAACATCCAGACACGCGCTCTCGCGTACCGCGAAACCGCTCGACAGTCCGCCCTC 844
Db 212 AspAlaGluThrGluSerThrLeuLeu-----SerGlyLeu 223
Qy 845 TACCAAAAATAACAGACGGGACGC-----ATCGACTGGCAGAGC 883
Db 224 GlyAspLysGlySerThrAlaGluPheIleAlaAlaSerSerAspValHisTrpValAsp 243
Qy 884 GTCCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAAA 943
Db 244 ---ArgValArgValIleSerAspProGluLys---ValArgGlyTyrPArgProArg 261
Qy 944 CCGCGCATTCGCGGAGGCTGCAAGCGCTCAACAGCCCGCAAAACCGCTCTATCTG 1003
Db 262 GlyTrpLeuMetLysGluLeuLeuProIleIleGlnSerAlaArgLysArgValGluIle 281
Qy 1004 GTTTCACCCCTATTTCGCTCCCTACAAAATCCGGCACACAGACGACGCTGGCAAAATGGTGCAG 1063

```
|||||
282 ValSerProTyrPheIleProGlyLysLysGlySerLysIleLeuGlyAspLeuValAsp 301
QY 1064 GACGGCATAGACGTTACCGCTCTGACCAACATCGCTACAGGGGACGACGAGTTGCGCGCGTC 1123
Db 302 AspGlyValGlnValAlaValLeuThrAsnSerLeuAlaIleThrAspValAlaVal 321
QY 1124 CATTCGCGCTAGCTCAATACCGAAACCGCTGCTCAAAAGCGGCGCATCAAACTCTACGAG 1183
Db 322 HisGlyAlaTyrAlaAsnTyrArgLysArgLeuLeuArgLysGlyValGlnLeuPheGlu 341
QY 1184 CTGCAACCAACCATCGCTCCCGCCACAAAGAGAAAGCGCTGACCGGAGCTCCGTA 1243
Db 342 LeuGlnProPheSerArgGlnProLysIle-----SerValPheGlySerLysGly 358
QY 1244 ACCAGCTCGCATGCCAAACCTTCATTGTGACGGCAACACGCATCTTCATCGGCTCATTC 1303
Db 359 AlaSerLeuHisThrLysAlaPheSerValAspAspArgIleGlyPheValGlySerPhe 378
QY 1304 AACCTCGACCCCGTTCGCGAGCTCAATACCGAAATGGCGTCTGTCATCGAAGCCGCC 1363
Db 379 AsnPheAspProArgSerValSerLeuAsnSerGluMetGlyValLeuPheGluAspGln 398
QY 1364 AAA---ATCGCAGACACATGAGCGCACCTCGCGGATACACACCGAATACGCTTAC 1420
Db 399 LysLeuValMetGluLeuArgHisArgPheLysSerGluThrAlaProGluAlaSerTyr 418
QY 1421 CGCGTTACCTCTGCACAAACACGCGCTCGCAATGGCAC-----GATCCGCGCCACCGA 1474
Db 419 ArgLeuGluLeu---LysAsnGlyValLeuArgTrpHisGlySerAspGluGlyArgPro 437
QY 1475 AAAACCTACCGAAGCAACCGGACCGCAAACTTTGGAAACGCATCGCGCAAAATCTCTA 1534
Db 438 GlnThrTyrAlaArgGluProGluAlaAlaLeuPheArgIleLeuSerAlaLeuVal 457
QY 1535 TCCCTGCTGCCATCGAGGTTTATTA 1561
Db 458 ArgTyrLeuProIleGluSerGlnLeu 466

RESULT 33
ID Q8UEX3 AGRT5 PRELIMINARY; PRT; 518 AA.
AC Q8UEX3; Q7CYU7;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DE 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Phospholipase D family protein (AGR_C_3007p).
GN OrderedLocusNames=AGR_C_3007, Atul6307
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Secubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Senhimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
```

```
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu P.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328 (2001).
DR EMBL; AE009120; AAL42632.1; -; Genomic_DNA.
DR EMBL; AE008085; AAK87408.1; -; Genomic_DNA.
DR PIR; AB27777; AB27777.
DR PIR; G97556; G97556.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 518 AA; 58096 MW; F08804A62B810AE2 CRC64;

Alignment Scores:
Pred. No.: 7.32e-37 Length: 518
Score: 807.00 Matches: 204
Percent Similarity: 54.6% Conservative: 85
Best Local Similarity: 28.8% Mismatches: 196
Query Match: 28.3% Indels: 44
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x Q8UEX3_AGR5 (1-518)
QY 53 CTCATTTCCCTTTTATGCTCTCTCTGTTTCATGTTCTTCATGTTGCCCCACTGGAA 112
Db 8 IleIleIleLeuLeuMetLeuIleGlyProSerLeuPheValIleGlyLysGlnArg 27
QY 113 GAACGCGAGAAAGCGGTCATTTCAATACTTCCAAACCTGTCCTCTCGACAAACATCTCTG 172
Db 28 GluLysAlaIleProLysArgProSerThrAlaLeuProValThrGluAspGluThrAla 47
QY 173 CAATTCGCGCACACCCCTCAT---AACACGGGTATCCGAC-----ATCATCTG 220
Db 48 LeuAspArgHisTrpGlnSerIleArgAsnGlyTrpAsnGluLysAsnAlaLeuCysLeu 67
QY 221 CTCGACGACCCCCACGAGACCTTTCGCGCGCGCGCTTCATCGAATTCGCGAACAC 280
Db 68 LeuHisSerAsnLeuAspAlaPheAlaValArgValAlaAlaAlaArgAlaGlyArg 87
QY 281 AGCTTCGATTTGCAATACATATTGGCGCAACGACATTTCCGCGAGCTCTGTTCAC 340
Db 88 SerLeuAspLeuMetTyrTyrMetTrpAsnAlaAspLeuThrGlyArgLeuMetMetArg 107
QY 341 CTCATGTACCTTGGCGGAGAACGGCGGTGCGGTACGCTGCTGCTGTTGGACGACAAAC 400
Db 108 GluValIleAlaAlaAlaAspArgGlyValArgValArgLeuLeuAspLeuGly 127
QY 401 ACGCGCGGTTGGACGATCTCTCTGCTCCCTCGACAGCCATCCCAATATCGAAGTCGC 460
Db 128 ValSerMetSerAspArgIlePheHisAlaIleAspSerHisProAsnIleGluLeuArg 147
QY 461 CTGTTCAACCCCTTCGTCCTACGCAAA-----TGGCGCGCACTCGGTACTCTGACC 511
Db 148 LeuPheAsnProThrArgAlaArgGluAsnIleLeuHisArgSerLeuGluLeuValLeu 167
QY 512 GACTTCCCCCGCTCAACCGCGGATGCACAAACAAATCTTTACCGCGACACCGCGCC 571
Db 168 ArgPheArgSerValAsnArgArgMetHisAsnLysAlaIleAlaAspGlyArgAla 187
QY 572 ACCATATCTCGCGGACGCAATATCGCGACGAATCTTCAAGTCCGTCGAGGACACCGCTT 631
Db 188 ValIleValGlyArgAsnIleGlyAspAlaTyrPheAspAlaAlaGluArgAlaAsn 207
QY 632 TTCGCGGACCTGGACATCTCCGCCACCGCGAGCGTGTGCGGAGAGTATCGCACACTTC 691
Db 208 PheHisAspPheAspIleLeuGlyPheGlyGlyIleValAlaAspAlaThrGluIlePhe 227
```

```
QY 692 GACCGCTACTGGCAAGGCAATTCGCCCCACAAAC-----GCCACG 730
DB 228 AspAspTyrTrp-----AsnSerAlaValSerValProValArgSerLeuLeuAlaArg 245
QY 731 GCATCATCCGAGCGGCAACATCCGGCAAGGCTCTCAAGCACTCGGATACAACGACGAA 790
DB 246 ArgProAsnLysLeuAlaLysLeuArgGluLeuLeuAspAlaLeuProGlnSerGlu--- 264
QY 791 ACATCCAGACACGCGCTCTCGCTACCGCAACCGTCGACAGCTCGCCCTCTACCAA 850
DB 265 -----AlaAlaArgProTyrLeuGluArgValGluSerGln-----TyrGly 278
QY 851 AAA-----ATACAGACGGAGCATCGACTGGCAGAGCGCTCCAAACCCGCG---CTG 898
DB 279 ArgAspHisPheLeuMetSerAspArgLeuHisTrp-----ValAspThrAlaAspVal 296
QY 899 ATCAGCGACACCCCTGCAAAAGGATCTCAGCGCGACCGCGCAAAACCCGCG-----ATT 952
DB 297 LeuAlaAspProProGluLysAlaAlaGlyLys---ArgArgLysGlyHisAsnPheLeu 315
QY 953 GCCGGAGGCTCAAGACGCGCTCAACACAGCCGCAAAAGGCTCTATCTGTTTCACCC 1012
DB 316 MetGluSerLeuLeuProLeuMetGlnAlaAlaGlyGluSerLeuHisIleThrSerPro 335
QY 1013 TATTTTCGCTTACAAAATCCGGCACAGACGACTGGCAAAACCTGGTCGAGCGGCATA 1072
DB 336 TyrPheIleProGlyLysGlnGlyValGluIlePheLeuAspLeuAlaGluArgGlyVal 355
QY 1073 GACGTTACGCTCTGACCAACTCGCTACAGCGGACCGGCTGGCGCGCTCAATTCGCGC 1132
DB 356 SerLeuAlaIleLeuThrAsnSerLeuAlaIleThrAspValAlaAlaValHisAlaGly 375
QY 1133 TAGGTCAAATCCGAAACCGCTGCTCAAGCGCGCATCAAACTCTAGAGCTCAACCC 1192
DB 376 TyrAlaArgTyrArgLysProLeuLeuSerGlyGlyValArgLeuHisGluLeuArg--- 394
QY 1193 AACCATGCGCTCCCGCCACAAAAGCAAAAGGC-----CTGACCGCGAGCTCCGTA 1243
DB 395 -----SerGlnAlaAspGlnGlySerPheThrLeuArgGlySerGlyGln 409
QY 1244 ACCAGCTGATGCAAAACCTTCAATTTGAGCGGCAAAACGCATCTTCATCGGTCATTC 1303
DB 410 AlaSerLeuHisThrLysAlaPheThrArgAspGlyGluThrGlyTyrIleGlySerLeu 429
QY 1304 AACCTCGACCCCGCTCCGACGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCC 1363
DB 430 AsnPheAspProArgSerAlaSerLeuAsnThrGluMetGlyValValPheAsnSerAla 449
QY 1364 AAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCAACACCCGAA---TAGCGCTAC 1420
DB 450 ProLeuValAlaArgMetAspGluIlePheAlaGluIleArgArgThrMetSerPhe 469
QY 1421 CGGTTACCTCGCAAAACACACCGCTGCAATGG-----CAGCATCCGCCACCCGGA 1474
DB 470 GluLeuAspIleAspSerAlaAsnArgIleValTrpMetThrGluGluArgGlyGlnPro 489
QY 1475 AAAACCTACCCGAGAACCGAACCAAACTTTGGAAACGCATCGCGCGAAATCCTA 1534
DB 490 LysIleTyrArgArgGluProAspAlaAlaIleSerArgArgIleAlaGlyIleMet 509
QY 1535 TCCCTGCTGCCCATCGAAGGTTTATTA 1561
DB 510 ArgValLeuProLeuGluSerGlnLeu 518
RESULT 34
Q984B3_RHILO
ID Q984B3_RHILO PRELIMINARY; PRT; 466 AA.
AC Q984B3
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cardiolipin synthetase family protein.
```

```
OrderedLocusNames=mlr8077;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OX Phyllobacteriaceae; Mesorhizobium.
RN NCBI_TaxID=381;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000112; BAB53717.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 466 AA; 52117 MW; 693301E5080AA788 CRC64;

Alignment Scores:
Pred. No.: 9,99e-37 Length: 466
Score: 804.50 Matches: 193
Percent Similarity: 53.9% Conservative: 61
Best Local Similarity: 41.0% Mismatches: 180
Query Match: 28.2% Indels: 37
DB: 2 Gaps: 10

US-10-665-990A-13 (1-1561) x Q984B3_RHILO (1-466)
QY 200 GGGCTATCCGACATCTACCTGCTCGACGACCCCGAACGACCTTGGCGCGCGCGCC 259
DB 16 GlyLysSerGlyLeuLeuLeuIleSerAspAsnTyrAspAlaPheAlaAlaArgValLeu 35
QY 260 CTTATCGAATCTGCGCAACACAGCTCGATTTGCAATCTACTACATTTGGCGCAACGACATT 319
DB 36 AlaAlaArgGlyAlaAlaArgThrLeuAspLeuMetTyrTyrLeuTrpHisAspAspHis 55
QY 320 TCGCGCAGGCTGCTGTTCAACCTCATCTGCTCCCGCAGAACCGCGGCTGGCGTACGC 379
DB 56 ThrGlyArgLeuLeuLeuGlnValValArgAlaAlaGlnArgGlyValArgValArg 75
QY 380 CTGCTGTGTGACGACAAACACCGCGGGTGTGGACGATCTCTCTGCTCGCCCTCGACAGC 439
DB 76 MetLeuLeuAspAspValAsnProArgLysSerAspAlaIleTyrLeuAlaLeuSerAsn 95
QY 440 CATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCTGCTACGCAAA-----TGG 490
DB 96 HisProAsnIleGluLeuLysLeuPheAsnProSerGlyIleArgAlaArgGlyLeuMet 115
QY 491 CGCGCATCTCGGCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGCGCAACAATCC 550
DB 116 ArgGlyAlaGluValLeuLeuArgLeuPheAlaLeuThrArgArgMetHisAsnLysAla 135
QY 551 TTTTACCGCCGACAAACCGCGCACCATCTCGCGGAGCAATATCGCGCAGCAATACTTTC 610
DB 136 TrpIleAlaAspAspAsnIleAlaIleValGlyGlyArgAsnValGlyAspAlaTyrPhe 155
QY 611 AAAGTCGGTGAGGACACCGGTTTTCGCGCAGCTCGACATCTCGCACCGCGCAGCGCTGTC 670
DB 156 AspAlaAlaGlu---ThrAsnPheArgAspLeuAspMetLeuLeuLeuGlyProAlaVal 174
QY 671 GCGGAAGTAGTCGACGACTTCGACCGGCTACTGGCAAGCCATTCCGCGCCCAACGCCACG 730
DB 175 GlnGlnThrAlaGlnIlePheGlnThrPheTrpValCysGlnAspAla----- 190
```

```
QY 731 CGCATCATCCGCGCGCAACATCGCGAAGGCTCTTCAAGCACTCGGATACACACGAA 790
Db 191 -----LysProfilAlaGluLeuGlyAlaAlaPro 201
QY 791 ACATCCAGACACGCGCTCGTGGCTACCGCGAAACCGTGAAGCACTGCGCCCTTACCAA 850
Db 202 GlySerHisAlaProTyrPheGluGlyArgGluGluLysThrGluSerThrLeuLeuSer 221
QY 851 AAAATACAG-----ACGGACGCGATCGACTGG 877
Db 222 GlyIleArgAspLysGlySerIleAlaGluPheIleSerAlaSerAsnValHisTrp 241
QY 878 CAGAGCGTCAAAACCGGCTGATCAGCAGCACCCCTGCAAAAGGACTCGACGCGACGC 937
Db 242 ValGlu---ArgValArgValIleSerAspProGluLys---ValArgGlyTrpArg 259
QY 938 CGCAACCGCCCATCGCGGAGGTGCAAGACGCGCTCAACACGCCCGAAAAAGCGTC 997
Db 260 ProArgSerTrpLeuMetLysGluLeuLeuProIleIleGlnSerAlaArgLysArgVal 279
QY 998 TATCTGGTTTCCCTATTTCCTCCTACAAAATCCGCGACAGCGCACTGCAAAACTG 1057
Db 280 GluIleValSerProTyrPheIleProGlyLysLysGlySerLysIleLeuGlyAspLeu 299
QY 1058 GTGCGAGGCGCATAGACGTTTACCGTCTGACCAACTCGCTACAGCGGACGACGTTGCC 1117
Db 300 ValAspAspGlyValGlnValAlaValLeuThrAsnSerLeuAlaAlaThrAspValAla 319
QY 1118 GCGGTCCATTCGCGTCAATACCGAAACCGCTGCTCAAAACCGGATCAAACTC 1177
Db 320 AlaValHisGlyAlaTyrAlaAsnTyrArgLysArgLeuLeuArgMetGlyValGlnLeu 339
QY 1178 TAGAGGTGCAACCCAAACATCGCGTCCCGGCCCAAAAAGACAAAGGCTGACCGGACG 1237
Db 340 PheGluLeuGlnProPheSerArgGlnProLysIle-----SerValPheGlySer 356
QY 1238 TCCGTAACGAGCTCGATGCCAAACCTTCATTGTGGAGCGGCAACGCACTTTCATCGGC 1297
Db 357 LysGlyAlaSerLeuHisThrLysAlaPheSerValAspAsnArgIleGlyPheValGly 376
QY 1298 TCATTCAACCTCGACCCCGTTCGCGCGCTCAATACCGAAATGGGGTTCGTCATCGAA 1357
Db 377 SerPheAsnPheAspProArgSerValSerLeuAsnSerGluMetGlyValLeuPheGlu 396
QY 1358 AGCCCCAAA---ATCGCAGAACAGATGGAGCGCACCTCGCGATACACACCGCAATAC 1414
Db 397 AspGluAsnLeuValAlaGluLeuArgHisArgPheLysSerGluIleAlaProGluAla 416
QY 1415 GCCTACCGGTACCTCGACAAACACACCGCTGCAATGGCAC-----GATCCCGCC 1468
Db 417 SerTyrArgLeuGluLeu---LysAsnGluValLeuArgTrpHisGlySerAspGluGly 435
QY 1469 ACCCGAAACCTTACCGCAACCGAAGCGCAAACTTTGAAACGGATCGCGCCAAA 1528
Db 436 ArgLeuGlnThrTyrThrArgGluProGluAlaAlaTrpPheArgArgIleLeuAlaAla 455
QY 1529 ATCTATACCTCGTCCGCAATCGAAGTTTATTA 1561
Db 456 LeuValArgHisLeuProIleGluSerGlnLeu 466
RESULT 35
Q6FB53 ACIAD PRELIMINARY; PRT; 520 AA.
AC Q6FB53;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative phospholipase D protein.
GN OrderedLocusNames=ACIAD1881;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
```

```
[1]
NUCLEOTIDE SEQUENCE.
RP STRAIN=ADP1;
RC PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
DR EMBL; CR543861; CAG68710.1; -; Genomic DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 60385 MW; 1DD312966EA749A0 CRC64;

Alignment Scores:
Pred. No.: 3.9e-36 Length: 520
Score: 794.00 Matches: 186
Percent Similarity: 53.1% Conservative: 76
Best Local Similarity: 37.7% Mismatches: 177
Query Match: 27.8% Indels: 54
Db: 2 Gaps: 11

US-10-665-990A-13 (1-1561) x Q6FB53_ACIAD (1-520)
QY 182 CACACCCCTCATAAC-----AACGGGCTATCCGACATCTACCTGCTCGACGACCCCCAC 235
Db 49 HisThrAspGluAsnLeuGlnLysGlyLeuThrAlaTyrIleProLeuAspAspAlaPhe 68
QY 236 GAAGCCCTTGGCGCGCGCGCCCTTATCGAAATCTGCGAACACAGCCTCGATTGCAA 295
Db 69 IleSerIleAlaSerArgIleTyrLeuIleArgHisAlaLysHisGlnIleAspLeuGln 88
QY 296 TACTATATTTGGCGAAGCAGCATTTCCGGGAGGTGCTGTTCACCTCACTGATCTGCC 355
Db 89 TyrTyrIleTrpHisAsnAspPheValGlyGlnLeuIleLeuAsnGluLeuLysAla 108
QY 356 GCAGAACGCGGTGCGGTACGCTGTGTGGACGACAAACACACGCGGGTGGAC 415
Db 109 AlaAspArgGlyValLysIleArgLeuLeuIleAspGlnAsnGlyThrArgLeuAsp 128
QY 416 GATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCAACCCCTTC 475
Db 129 SerThrPheGlnAlaLeuGlnHisProLeuPheGluIleLysValPheAsnProTyr 148
QY 476 GTCCTACGCAATGGCGGCACTCGGTACTCGTACCGGACTTCCCGCGCTCAACCGCGC 535
Db 149 LysPheArgLysLeuArgPheValAspPheLeuPheArgProAsnGlnIleAsnHisArg 168
QY 536 ATGCACACAAATCCTTTTACCGCGACAAACCGCGCCACCATCTCGGCGGACGCAATATC 595
Db 169 MethHisAsnLysLeuIleAlaAspAsnThrIleAlaValThrGlyValArgAsnIle 188
QY 596 GCGCAGCAATACTTCAAAAGTCGGTGGAGACACCGTTCGCGGACCTCGACATCTCGCC 655
Db 189 SerSerGluTyrPheAspAlaSerTyrGlnPheGlnPheThrAspLeuAspIleLeuPhe 208
QY 656 ACCGCGCGCTGTCGGGCAAGTATCGCACACTTCGACCTCGACCGCTACTGGGCAAGCATTC 715
Db 209 TyrGlyThrAlaAlaSerGlnAlaGluLysValPheGlnThrPheTrpAspSerProLeu 228
QY 716 GCCCAACACGCGCAGCATCTCGCAGCGGCAACATCGGCAAGGGTCTTCAACGACCTC 775
Db 229 SerValThrThrGlnLysLeuIleThrLysGlyGlnProAsnAspLeuGlnLeuArg 248
QY 776 GGATACAAACGACGAAACATCCAGACACCGCGTCTCGCGTACCGCGAAACCGTCGAACAG 835
Db 249 Ser-----LysTyrLysGlu---IleGluArg 256
```

```
QY 836 -----TCGCCCCCTCTACCAAAAATAACAGACGGGACGCGATCGAC----- 874
DB 257 ValAspThrProThrgluAspLysIleAsnGluAlaGlnIleGluLeuAsnGluAlaLeu 276
QY 875 -----TGGCAGAGCTCCAAACCCCGCTGATCAGGACGACCCCT 913
DB 277 LysGlnArgTyrSerValLysTrp-----AlaLysAlaHisPheValAlaAspSerPro 294
QY 914 GCAAAAGAGCTCGACCGCGCGCGCAACCGCGATTCGCGGAGGCTCAAGACGCG 973
DB 295 LysLysIleHisGlyHisAlaIleGlyAsnGluLeuIleTyrAsnGlnMetPheSerIle 314
QY 974 CTCAAACAGCCGCAAAAAGCGTCTATCGTGTTCACCTATTTCGCTCCCTACAAAATCC 1033
DB 315 MetGlyHisProGlnLysHisLeuGluLeuValSerAlaTyrPheValProThrGlnAla 334
QY 1034 GGCACAGACGACTGGCAAACTGGTGACGACGCGATAGACGTTACGCTCAGCAAC 1093
DB 335 GlyThrGlnTyrLeuSerAsnLeuSerGlnGlnAspValLysIleArgValLeuThrAsn 354
QY 1094 TCGCTACAGGCGACCGAGTTCGCGCTCCATTCGCGTACGTCAAATACCGAAACCG 1153
DB 355 SerPheAlaAlaAsnAspValAlaIleValHisAlaPheTyrSerGlnTyrArgLysGlu 374
QY 1154 CTGCTAAAGCCGCGCATCAACTCTACGAGCTGCAACCC----- 1192
DB 375 LeuIleLysAsnGlyIleGlnLeuTyrGluPheLysProIleIleAspArgLysGluPro 394
QY 1193 -----AACATGCGCGTCCCGCCGCAAAAGACAAAGCGCTGACC 1231
DB 395 ThrTrpTyrGluLysMetThrGlyArgValIleProAla-----Lys 408
QY 1232 GGCAGCTCCGTAACAGCGCTGCATGCCAAACCTTCATTGFGAGCGCAACGCGATCTC 1291
DB 409 GlyLysLysSerSerLeuHisAlaLysPhePheAspValAspGlyLys---ValPhe 427
QY 1292 ATCGGCTCATCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGTCTC 1351
DB 428 IleGlySerPheAsnPheAspProArgSerAlaTyrLeuAsnThrGluValGlyLeuVal 447
QY 1352 ATCGAAAGCCCAAAATCGCAGAACACATGAGGCGCACCTCGCGCATACACACCGGAA 1411
DB 448 ValGluSerGluAspLeuGlnAsnGluIleThrLysValLeuAspGluTyrLeuProThr 467
QY 1412 TAGCCCTACCGGCTTACCTCGACAAACACACCGCTGCAATGGCAGCAT---CCGCGC 1468
DB 468 IleAlaTyrGlnLeuLysLeuAspLysAsnGlyAsnLeuValTrpLeuAspHisAsnAla 487
QY 1469 ACCGCAAAACCC-----TACCCGAAACGCAACCCGCAACCAACTTTGGAAACGATCGCC 1522
DB 488 AspGlyGlnThrIleGluTyrHisHisAspProGluThrThrArgPheGlnArgPheMet 507
QY 1523 GCAAAATCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
DB 508 MetLysAlaValSerTyrLeuProIleGluTrpMetMet 520

RESULT 36
Q9HTP4_PSEAE
ID Q9HTP4_PSEAE PRELIMINARY; PRT; 529 AA.
AC Q9HTP4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=PA5310;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
```

```
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentini E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004943; AAG08695.1; -; Genomic_DNA.
DR PIR; F82983; F82983.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; FLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 529 AA; 59199 MW; DABF5ED5881E63DE CRC64;

Alignment Scores:
Pred. No.: 5,05e-36 Length: 529
Score: 792.00 Matches: 200
Percent Similarity: 52.5% Conservative: 86
Best Local Similarity: 36.7% Mismatches: 189
Query Match: 27.8% Indels: 70
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x Q9HTP4_PSEAE (1-529)
QY 47 CGCAGCGCTATTCCTCTTATCGCTCTCTCTGTTCTCATGTTCTTCATGTTGCCCGCA 106
DB 15 ArgArgPheLeuLeuLeuAlaLeuLeuAlaLeuLeuAlaLeuLeuAlaLeuLeuAlaLeu 34
QY 107 -----CTGGAAGAACGCGAGGAAAGCGGT 130
DB 35 GlnProSerSerAlaLeuProAlaGluGlyThrTrpLeuAlaArgGlnAlaGlu----- 52
QY 131 CATTCATTAATCTTCCAAACCTGCTCTCTGGCAACATCTCTGAAATCCGGCACACCCCT 190
DB 53 -----IleGlnGlyArgAspHisPro 59
QY 191 CATAACACGCGGTATCCGACATCTACCTCTCGACGACCCCGACAGCCCTTGGCGC 250
DB 60 -----GlyGlnSerGlyPheHisLeuLeuSerAlaSerGluAspAlaPheValAla 76
QY 251 CGCGCGCGCTTATCAATCTGCGAACAACAGCTCGATTGCAATACTACATTTGGCGC 310
DB 77 ArgAlaAlaLeuIleArgAlaAlaGlnArgSerLeuAspIleGlnTyrTyrIleValHis 96
QY 311 AACGACATTTCCGGCAGGTGCTGTTCAACCTCATGATCTTCCCGCAGAACCGCGGTG 370
DB 97 AspGlyLeuThrThrArgAlaLeuAlaTyrGluLeuLysAlaAlaAspArgGlyVal 116
QY 371 CGCGTAGCGCTGCTTGGACGACACACACGCGGGTGGACCATCTCTCTGCTCGCC 430
DB 117 ArgValAlaLeuIleAspSerThrAlaSerAspGlyTrpAspTyrGluIleGlyVal 136
QY 431 CTGACAGCGCATCCCAATATCGAAGTGGCGCTCTTCAACCCCTTCGCTTACGCAAA--- 487
DB 137 LeuSerAlaHisProAsnIleGlnValArgLeuPheAsnProLeuHisLeuGlyArgAla 156
QY 488 -----TGGCGCGCATCTCGGCTACCTGACCGATCTCCCGCGCTCAACCCCGCATG 538
DB 157 ThrGlyIleThrArgGlyValGlyArgLeuPheAsnLeuSerGlnGlnHisArgMet 176
QY 539 CACAACAAATCTTTTACCGCGCACACCGCGCCACCATCTCGCGCGGAGCATATCGGC 598
DB 177 HisAsnLeuLeuTrpLeuAlaAspGlyThrAlaAlaIleValGlyArgAsnLeuGly 196
QY 599 GACGAATACTTCAAAGTCGGTGAGGACACCGTTTTTCGCGCGACCTGGACATCTCTCGCCACC 658
```

Db	197	AspGluTyrPheAsnAlaLysProGluMetAsnPheThrAspLeuAspLeuLeuGlyVal	216
Qy	659	GGCAGCGCTCGTGGCGGAAGTATCCACGACTTCGACCGCTACTGGGCAACCCATTCGCGC	718
Db	217	GlyProIleAlaAsnGlnLeuSerHisSerPheAspGlnTyrTrpAsnSerAlaIleSer	236
Qy	719	CACAACGCCACG-----CGCATCATCCGACGCGCAACATCGCAGCAAGGTCTT	766
Db	237	ArgProIleGluAspPheLeuTrpArgAlaProTyrProGlyGluLeu-----	252
Qy	767	CAAGCATCTCGGATCAACGAGCAACATCCAGACCGCGCTCTCGCGCTAC-----CGC	820
Db	253	-----GluSerAlaArgLysLeuGlnArgTyrLeuArgLys	265
Qy	821	GAACACCGTCGAACAGTCG-----CCCTCTCTACCAAAAAATACAGACGGGACGC	868
Db	266	GluSerValLysGluSerGlyTyrIleArgHisLeuPheAspArgGlyAspGlnProArg	285
Qy	869	ATC-----GACTGGCAGAGC-----GTCCAAACCCCGCTGATCAACGACACCCCT	913
Db	286	LeuGlyAsnTrpLeuGluAsnLeuThrTrpAlaArgAlaGluAlaIleTrpAspAlaPro	305
Qy	914	GCAAAAGGACTGCAGCCGCGCGCGCAACACGCCG-----ATTGCCGGAGGCTGC	967
Db	306	LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuLeuSerProHisLeuAla	325
Qy	968	GACGCGCTCAACAGCCCGAAAAAGCGTCTATCTGCTTTTCCACCTATTTCGTCCTACA	1027
Db	326	GlyLeuPheLysGlyValGlnLysGluLeuLeuValSerAlaTyrPheValProAla	345
Qy	1028	AAATCCGGCCACAGCAGCAGCTGGCAAACTGGTGACGCGCATAGACGTTACGTCCTG	1087
Db	346	LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu	365
Qy	1088	ACCAACTCGCTACAGGCGACGCGTTCGCGCGCTCCATTCGCGCTACGTCAATACCGA	1147
Db	366	ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg	385
Qy	1148	AAACCGCTGCTCAAGCCGCGCATCAAACTCTACGAGCTGCACCCACCATGCGTCCCC	1207
Db	386	MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro	405
Qy	1208	GCCACA---AAAGACAAAGGCGCTGACCGCGAGCTCCGTAACACGCTGCATGCCAAAC	1264
Db	406	LeuSerGlyAlaProTrpArgLeuHisGlySerSerSerAlaSerLeuHisSerLysAla	425
Qy	1265	TTCATTGTGGACGCCAAACGCACTTCATCGGCTCATTCACCTCGACCCCGTTCGCGA	1324
Db	426	MetValPheAspArgArgLysValPheIleGlySerPheAsnPheAspProArgSerIle	445
Qy	1325	CGGCTCATACCGAAATGGCGTCGTATCGAAGGCCCAAAATCGCAGAACAGATCGAG	1384
Db	446	LeuTrpAsnTrpGluValGlyValIleValAspSerProLeuLeuAlaGluGlnValArg	465
Qy	1385	CGC---ACCCTCGCCGATACACACACCCGAAATACGCTACGCGTTACCTCGACAAACAC	1441
Db	466	GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer	485
Qy	1442	AAC-----CGCTGCAATGGCAGAT-----CCCGCCACCCGAAAACTACCCG	1486
Db	486	GlySerArgProLysLeuValTrpIleAspGluArgAspGlyArgAlaGlnValLeuArg	505
Qy	1487	AACGAACCCGAGCCAAACTTTGAAACGCAATCCCGCCAAATACTCTATCCCTGCTGCC	1546
Db	506	HisGlnPro---GlySerLeuTrpArgLeuAsnAlaTrpValAlaGlyMetIleGly	524
Qy	1547	ATCGAAGGTTTATTA	1561
Db	525	LeuGluLysMetLeu	529

```
QY 464 TTCACCCCTTGGTCTCTACGCAAAATGG-----CGCGCACTCGGTACTCGTACC 511
Db 163 PheAsnProPheGlyAlaProArgLeuGlyMetPheAlaA:GThrAlaAsnValPheThr 182
QY 512 GACITTCCTCCGCTCAACCGCGCATGACACAAATCCTTTACCGCGCAGAACCGGCC 571
Db 183 ArgIleAspAsnPheThrArgArgMetHisAsnLysAlaMetIleSerAspAsnGlnIle 202
QY 572 ACCATACCTCGCGGACGCAATATCGCGCAGCAATACTTCAAAGTCGGTGAGGACACCGCTT 631
Db 203 AlaIleValGlyGlyArgAsnLeuGlyAspGluTyrPheAsnAlaSerProThrLeuGln 222
QY 632 TTCGCCGACCTTGACATCTCTGCACCGCGAGCGTGTGCGGAGTATTCACGACTTC 691
Db 223 PheArgAspLeuAspValLeuAlaAlaGlyProValThrArgAlaValSerAlaSerPhe 242
QY 692 GACCGCTACTGGGCAAGCATTCGCGCCACACAGCCGCGCATCATCCGCGAGGCGAAC 751
Db 243 AspAlaIleTyrPheSerAlaLeuThrTyrProLeuProAlaLeuAsnArgArgTyr 262
QY 752 ATCGGCAAGGTCTTCAAGCACTCGGATACACAGCAAAACATCCAGCAGCGCTCTCTG 811
Db 263 AspAlaIleAspLeuAspAla-----AlaArgAspAlaLeuArg 275
QY 812 CGCTAC-----CGCGAAACCGTCGACAGTCCGCCCTC 844
Db 276 AlaHisTyrAlaAlaAsnAlaThrProTyrAsnAlaLysProLeuAsnAlaThrProLeu 295
QY 845 TACCAGAAATACAGCGGAGCATCGAC-----TGGCAGAGCGTCCAAACCGCGCTG 898
Db 296 AlaAlaGlnIleAlaArgAsnGluLeuGlyLeuValTyrAlaSerAlaGlu-----Phe 313
QY 899 ATCAGCAGCACCCCTCGCAAAA---GGACTCGACCGCGACCGCCGCAACCGCGCATTCGCC 955
Db 314 ThrAlaAspSerProGluLysIleAlaAlaProAspAspSerTyrLysSerProProMet 333
QY 956 GGGAGGTGCAAGCGCGCTCAACACGCGCAAAACGCGTCTATCTGTTTTCACCCCTAT 1015
Db 334 GlnArgLeuPheAlaLeuThrArgAspAlaGlnArgGluPheLeuValLeuSerProTyr 353
QY 1016 TTCCTGCTTCAAAATCCGCGACAGCGCTGCAAAATCGGTGCGAGGACGGATAGAC 1075
Db 354 PheValProHisAspAlaGlyValAsnAlaLeuGlyArgLeuThrAlaArgGlyValArg 373
QY 1076 GTTACCGTCTGACCAATCGCTACAGCGCGACCGCTGCGCGCTCCATTCCGGGTAC 1135
Db 374 ValAlaIleLeuThrAsnSerLeuAlaAlaThrAspAlaIleAlaValGlnAlaGlyTyr 393
QY 1136 GTCAAATACGAAACCGCTGCTCAACGCGGCAATCAACTCTACGAGTGCACCCCAAC 1195
Db 394 AlaProTyrArgValProMetLeuGluArgGlyValGluLeuTyrGluTyrLysProAsp 413
QY 1196 CATGCCGTCGCCGCCCAAAAGACAAAGGCTGACCGGCAGC---TCGTAACCGAGCTG 1252
Db 414 -----ProGlyArgSerArgIleGlyMetLeuGlySerArgSerArgAlaSerLeu 430
QY 1253 CATGCCAAAACCTTCATTGTGGAGCGCAACGCGCATCTTCATCGCTCATTCACCTCGAC 1312
Db 431 HisAlaLysAlaTyrValIleAspArgLysIleLeuValIleGlySerMetAsnLeuAsp 450
QY 1313 CCCCGTTCGCGACGGCTCAATACCGAAATGGCGTGTCTATCGAAAGCCCAAAATCGCA 1372
Db 451 ProArgSerAlaHisLeuAsnThrGluLeuAlaLeuValIleHisSerProArgLeuAla 470
QY 1373 GAACAGATGGCGCAGCTCGCGATACACACA---CCGAAATACGCTACCGGTACC 1429
Db 471 AsnGluValAlaAlaAsnLeuPheAspGluValThrLysProThrIleSerTyrArgValThr 490
QY 1430 CTCGACAAACACAAAC-----CGC 1447
Db 491 LeuAlaProAspThrProGlyAlaAlaGlnThrThrGlyAlaGlyAlaProAlaTyrPro 510
QY 1448 CTGCAATGGCAGCATCCCGCCACCCGA-----AAACCTACCCGGAACGACCCGAGCC 1501
```

```
Db 511 LeuValTyrThrGluIleAlaAspGlyGlnValArgThrTyrSerValAspProAsnAla 530
QY 1502 AAACITTTGGAAACGCATCGCGCAAAATCCTATCTCTGCTGCCATCGAAGGTTTATTA 1561
Db 531 GlyPheTyrArgAsnLeuLeuThrGlyLeuCysLeuLeuLeuProIleAspGlnLeu 550
RESULT 38
Q88CA5_PSEPK
ID Q88CA5_PSEPK PRELIMINARY; PRT: 517 AA.
AC Q88CA5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=PP5276;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
[1] RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohelsel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016793; AAN70841.1; -; Genomic_DNA.
DR TIGR; PP5276; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
DR Complete proteome.
SQ SEQUENCE 517 AA; 57574 MW; 107E07900AA35751 CRC64;

Alignment Scores:
Pred. No.: 5,79e-35 Length: 517
Score: 773.00 Matches: 199
Percent Similarity: 53.6% Conservative: 83
Best Local Similarity: 37.8% Mismatches: 210
Query Match: 27.1% Indels: 34
DB: Gaps: 15

US-10-665-990A-13 (1-1561) x Q88CA5_PSEPK (1-517)
QY 47 CGCAGCGCTCATTCCTTTTATGCTCCTCTC-----TGTTTCATGTTCTTCATGCTG 100
Db 5 ArgAlaLeuProLeuLeuLeuValLeuLeuGlyValAlaGlyCysThrSerIleSer 24
QY 101 CCCCACTGGAAGAACGAGCGGAAAGCGTCATTTCAATACTTCCAAACCTGCTCCTCTG 160
Db 25 AlaPro-----ArgGluThrSerGlnAlaLeuProAlaHisGluSerAlaPheGly 41
QY 161 GACAACATCTGCAAAATCCGCGCACACCCCTCATCAACACGCGCTATCCGACATCTACTCT 220
Db 42 ArgSerValLeu-----ArgGlnAlaAlaProTyrGlyArgSerGlyPheArgLeu 59
QY 221 CTCAGCAGCCCCCAGAACGCTTGGCGCGCGCCCTTATCGAAATCTGCCGACAC 280
Db 60 LeuProAsnSerAsnGluAlaPheArgAlaArgAlaGluLeuIleArgAsnAlaGlnAla 79
```

```
QY 281 AGCCTCGATTTCATATACATTTGGCGCGCAAGACATTTCCGGCAGCGCTGCTGTTCAAC 340
Db 80 SerIleAspLeuGlnTyrTyrIleValHisAspGlyLeuSerThrArgAlaLeuValHis 99
QY 341 CTCATGTACCTTGGCGGAGACCGCGCGTGGCGTACGCTGCTGGAGCAACAAC 400
Db 100 GluLeuLeuArgAlaAlaAspArgGlyValArgValArgIleLeuLeuAspAspThrThr 119
QY 401 ACGCGCGGTGGAGCATCTCCTGCTGCCCTCGACAGCCATCCCAATATCGAAGTGGC 460
Db 120 SerAspGlyLeuAspThrValMetGlyThrLeuAspAlaHisProAsnIleHisIleArg 139
QY 461 CTGTTCAACCTTCTGCTCTACGAAA-----TGGCGGCGACTGGCTACCTG 508
Db 140 ValPheAsnProLeuHisLeuGlyArgSerThrGlyValThrArgAlaValGlyArgLeu 159
QY 509 ACCGACTTCCCGCGCTCAACCGCGCATGCACAACAATCTTTACCGCGCAACCGC 568
Db 160 PheAsnLeuSerArgGlnHisArgMetHisAsnLysLeuPheLeuValAspAsnSer 179
QY 569 GCCACCATACTCGCGGAGCGCAATATCGCGCAGCAATATCTCAAGTCGGTGAGCACAC 628
Db 180 MetAlaIleValGlyArgAsnLeuGlyAspGluTyrPheAspAlaGluProAsnLeu 199
QY 629 GTTTTCGGCGACCTGGACATCTCGCCACCGCGCAGCGTGTGGCGAAGTATCGCACAC 688
Db 200 AsnPheThrAspIleAspLeuLeuGlyValGlyProValAlaGluGlnLeuGlyHisSer 219
QY 689 TTCGACGCTACTGGCGAAGCATTTCCGCCCAACACGCGCATCATCGCAGCGGC 748
Db 220 PheAspGlnTyrTrpAsnSerAlaLeuSerArgProIleThrAspPheLeuTrpHisAsp 239
QY 749 AACATCGCGAAGGCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACCGCTC 808
Db 240 ProAspAlaAsnAspLeuArgAlaSerArgGlnArgLeuGluValSer-----Leu 256
QY 809 CTGCGCTACCGGAAACCGTCAACAGTCGCGCCCTCTACCAAAAAATACAGACG----- 862
Db 257 AlaLysAlaArg-----ThrGlnArgLysAlaLeuTyrAspArgLeuMetAlaTyrGln 274
QY 863 -----GGACGCTATCGAC---TGGCAGACG-----GTCCAAACCGCGCTGATC 901
Db 275 SerGlnProArgLeuAspValTrpArgAsnGluLeuIleTrpAlaHisAlaGlnAlaLeu 294
QY 902 AGCGACACCTTCGAAAGGACTCGACCGCGCCGCGCAACCGCG-----ATTGCC 955
Db 295 TrpAspAlaProSerLysValLeuAlaGluAspGluProAspProGlnLeuLeuSer 314
QY 956 GGGAGCTGCAAGCGCTCAACAGCCCGGAAAAAGGCTCTATCTGTTTCACCTAT 1015
Db 315 GlnGlnLeuAlaProAspLeuAlaAsnValHisArgGluLeuValLeuAlaSerAlaTyr 334
QY 1016 TTCGCTCCCTACAAATTCGGCACAGCGCATCGGCAAAACTGGTGCAGGACGGCATAGAC 1075
Db 335 PheValProGlyGluProGlyLeuLeuTyrLeuThrGlyArgAlaAspAlaGlyValSer 354
QY 1076 GTTACCGCTCCTGACCAACTCGCTACAGCGCACGCGTTCGCGCGCTCATTCGCGCTAC 1135
Db 355 ValLysLeuLeuThrAsnSerLeuGluAlaThrAspValProAlaValHisGlyGlyTyr 374
QY 1136 GTCAAATACCGAAACCGGTCTCAAGCGCGCATCAAACTCTACGAGTGAACCCAC 1195
Db 375 AlaProTyrArgAlaLeuLeuGluHisGlyValGlnLeuTyrGluLeuArgGln 394
QY 1196 CATCGCTCCCGCCACAAAACAGACAAAGCGCTGACCGGAGCTCCGTTACACGCTGCAT 1255
Db 395 ProGlyAspProSerAlaGlyArgLeuSerPheArgGlySerSerAspSerLeuHis 414
QY 1256 GCCAAACCTTCATTGTGGACGGCAAAACGATCTTCATCGGCTCATTCACCTCGACCC 1315
Db 415 ThrLysAlaIleValPheAspArgLysThrPheIleGlySerPheAsnPheAspPro 434
```

```
QY 1316 CGTTCCGACGGCTCAATACCAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAGAA 1375
Db 435 ArgSerValLeuTrpAsnThrGluValGlyValLeuValAspSerProGluLeuAlaGlu 454
QY 1376 CAGATGGAGCGCACCTCGCGCAT-----ACCACACCGCAATACGCTACCGCTTACC 1429
Db 455 ---TyrThrArgGluLeuAlaGlnGlnGlyMetAlaProAlaLeuSerTyrGlnValLys 473
QY 1430 CTCGACAAACACACCGCTCGCAATGG-----CAGATCCCGCACCGCAAAACCTAC 1483
Db 474 Leu---IleGlyAsnLysLeuValTrpAlaThrGluAspAsnGlyGlnArgHisValLeu 492
QY 1484 CCGAACGACCGCAAGCCAAACTTTGGAACGCGATCGCGCAAAATCCTATCCCTGCTG 1543
Db 493 ThrSerGluPro---GlyGlyIleTrpArgArgPheAsnAlaTrpIleSerLysAlaVal 511
QY 1544 CCCATCGAAGGTTTATTA 1561
Db 512 GlyLeuGluLysMetLeu 517

RESULT 39
Q4LJD9_9BURK PRELIMINARY; PRT; 540 AA.
AC Q4LJD9.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Phospholipase D/Transphosphatidyase precursor.
GN ORFNames=Bcen2424DRAFT_1235;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_taxid=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Irani S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHL0100070; EAM16198.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 38 Potential.
Alignment Scores: 6.2e-35 Length: 540
Pred. No.: 772.50 Matches: 198
Score: 52.2% Conservative: 83
Percent Similarity: 52.2% Mismatches: 208
Best Local Similarity: 36.8% Indels: 49
Query Match: 27.1% Gaps: 18
DB: 2
US-10-665-990A-13 (1-1561) x Q4LJD9_9BURK (1-540)
QY 41 AAAACACGACGCTCATTCCTTTATGCTCCTCTCTGTTTCATGTTCTTCATGGTTG 100
Db 21 GlnAlaArgAlaLeuValCysAlaLeuLeuProLeuAlaAlaCysAlaThrHisPro 40
QY 101 CCCCACTGGAA---GAACGAGCGGAAAGCGCTCATTTTCAATACTTCCAACCTGCTCTC 157
```



```

Db      41  ProAlaThrSerLeuAspArgProValSerHisAlaLeuSerAlaGluThrAlaThrPro 60
      |||  :|||  |||  :|||
Qy      158  CTGGACAATCTCTGCAATCCGGCACACCCCTCATATAACAAGGGCTATCCGACATCTAC 217
      |||  :|||  |||  :|||
Db      61  LeuArgAspAlaLeuAlaAlaProGluAlaAlaHis---ProGlyGlnSerGlyPheArg 79
      |||  :|||  |||  :|||
Qy      218  CTGCTCGACGACCCGCCAGACGCTTGGCGCCCGCGCCCTTATCGAATCTGCCGAA 277
      |||  :|||  |||  :|||
Db      80  LeuLeuAlaAspGlyAlaThrAlaLeuGlnMetArgIleAlaLeuAlaArgAlaAlaThr 99
      |||  :|||  |||  :|||
Qy      278  CACAGCTCGATTTCGAATACATTTGGCGCAACGACATTTCCGGAGGCTGTGTC 337
      |||  :|||  |||  :|||
Db      100  LysThrLeuAspMetGlnTyrtileAlaThrGluAspThrThrGlyLysLeuLeu 119
      |||  :|||  |||  :|||
Qy      338  AACCTCATGTACCTTGGCGCAGACGCGCGTGGCGGTACGCGTCTTGGACGACAA 397
      |||  :|||  |||  :|||
Db      120  GlyAlaAlaLeuTyrtileAlaAlaAspArgGlyValArgValArgMetLeuValAspLeu 139
      |||  :|||  |||  :|||
Qy      398  AACACGCGCGGTGGAGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAGTG 457
      |||  :|||  |||  :|||
Db      140  AsnPheHisAspIleAspArgValMetAlaAlaLeuAsnThrHisGlnAsnIleGluIle 159
      |||  :|||  |||  :|||
Qy      458  CGCTGTTCAACCCCTTCTGCTCTACGCAA-----TGGCGCGCATCGGCTAC 505
      |||  :|||  |||  :|||
Db      160  ArgValPheAsnProPheGlyAlaSerGlnArgGlyMetMetGluArgThrAlaAsnPhe 179
      |||  :|||  |||  :|||
Qy      506  CTGACCGATCTCCCGCTCAACCGCGCATGACACAAATCTTTACCGCGCAGAAC 565
      |||  :|||  |||  :|||
Db      180  PheThrArgIleAspSerPheThrArgArgMetHisAsnLysAlaMetIleAlaAspAsn 199
      |||  :|||  |||  :|||
Qy      566  CGCGCACCATCTACTCGCGGACGCAATATCGCGCACGCAATACTTCAAAAGTCGGTGAGAC 625
      |||  :|||  |||  :|||
Db      200  GlnLeuAlaIleValGlyArgAsnLeuGlyAspGluTyrtPheSerAlaSerProThr 219
      |||  :|||  |||  :|||
Qy      626  ACCGTTTTCGGCGACCTGGACATCTCTCCACCGGACGCGTGTGCGGAAAGTATCGCAC 685
      |||  :|||  |||  :|||
Db      220  LeuGlnPheArgAspLeuAspValLeuAlaAlaGlyProValThrSerAspIleSerLys 239
      |||  :|||  |||  :|||
Qy      686  GACTTTCGACCGCTACTGGCGAAGCATCTCCGCCACACGCGCGCATCATCGGACG 745
      |||  :|||  |||  :|||
Db      240  SerPheAspAspTyrtPalaSerAlaSerSerTyrt---ProLeuArgValLeuAsnHis 258
      |||  :|||  |||  :|||
Qy      746  GGCAACATCGGC---AAGGGTCTTCAAGCACTCGGATACACGACGAAACATCCAGAC 802
      |||  :|||  |||  :|||
Db      259  GlnThrPheAspProLysAspLeuAspAlaMet-----ArgAspGluLeuArgAspHis 276
      |||  :|||  |||  :|||
Qy      803  GCGTCTCTGCGCTACCGCGAAACCGTCGAA-----CAGTCG 838
      |||  :|||  |||  :|||
Db      277  -----TtpArgLysAsnAlaAspProTyrtAsnAlaLysProLeuAsnAlaThr 292
      |||  :|||  |||  :|||
Qy      839  CCCCTCTACCAAAAATACACAGCGGACGATCGAC-----TGGCAGAGGCTCAACAC 892
      |||  :|||  |||  :|||
Db      293  ProLeuAlaArgGlnIleAlaArgAspGluLeuValThrPalaProAlaGluPhe 312
      |||  :|||  |||  :|||
Qy      893  CGCTGTATCAGCAGACACCCCTGCAAAAGGACTCGACCGCGACCCGCG----- 940
      |||  :|||  |||  :|||
Db      313  LysVal-----AspAlaPro-----AspLysValAlaArgProThrGlyThr 326
      |||  :|||  |||  :|||
Qy      941  -----AAACCGCGATTCGCGGAGGCTGCAAGACGCGTCAACACGCGCGGAAAAAGC 994
      |||  :|||  |||  :|||
Db      327  TyrValSerProProMetGln---ArgLeuAlaGluLeuThrArgGlyAlaGlnLysGlu 345
      |||  :|||  |||  :|||
Qy      995  GTCTATCTGGTTTCCACCTATTTCTGCTCCCTACAAAATCCGCGACAGCAGCATCGCAAAA 1054
      |||  :|||  |||  :|||
Db      346  PheLeuAlaPheSerProTyrtPheValProHisAspAlaGlyValLysIleLeuGlyAsp 365
      |||  :|||  |||  :|||
Qy      1055  CTGCTGAGGAGCGCATAGACTTACCGTCTGTGACCAACTCGCTACAGGCGACGAGCTT 1114
      |||  :|||  |||  :|||
Db      366  ThrValAlaArgGlyValArgValAlaIleValThrAsnSerLeuAlaIlaThrAspAla 385
      |||  :|||  |||  :|||
Qy      1115  GCCCGCTCCATTCGCGCTACGTCAAAATCCGAAAAACCGCTGTCTCAAGCGCGCATCAAA 1174
      |||  :|||  |||  :|||

```

Alignment Scores:
 Pred. No.: 9,73e-35 Length: 542
 Score: 769,00 Matches: 200

```

Db      386  ValAlaValGlnSerGlyTyrtAlaProTyrtArgValProLeuLeuGlnArgGlyValGlu 405
      |||  :|||  |||  :|||
Qy      1175  CTCTACGAGCTGCACCAACCAACCATGCGTCCCGCCCAACAAAGAGAGGCTGACCGGC 1234
      |||  :|||  |||  :|||
Db      406  LeuTyrtGluPheLys-----SerGlnProAspGlnGlnProAlaArgLeuPheGly 422
      |||  :|||  |||  :|||
Qy      1235  AGC---TCCGTAACGAGCTGCATGCGCAAAACCTTCTTGTGGACGGCAACGCAATCTTC 1291
      |||  :|||  |||  :|||
Db      423  SerArgSerArgAlaSerLeuHisAlaLysAlaTyrtValIleAspArgGlnIleLeuVal 442
      |||  :|||  |||  :|||
Qy      1292  ATCGGCTCATTCACCTCGACCCCGCTTCCGACGGTCAATACCGAAATGGCGTCTGTC 1351
      |||  :|||  |||  :|||
Db      443  IleGlySerLeuAsnLeuAspProArgSerAlaHisLeuAsnThrGluLeuAlaLeuVal 462
      |||  :|||  |||  :|||
Qy      1352  ATCGAAAGCCCCAAAATCGCGAAGACAGATGGAGCGCCCTCGCGGATACCAACA---CCC 1408
      |||  :|||  |||  :|||
Db      463  IleHisSerProValLeuAlaGlnGlnAlaIlePheAlaArgValThrGlnPro 482
      |||  :|||  |||  :|||
Qy      1409  GAATACGCTACCGCTTACCTCGACAAACACAAAC-----CGCTCGCAA 1453
      |||  :|||  |||  :|||
Db      483  AspGluSerTyrtArgValThrLeuAlaLysArgThrAspGlyGlyProProAlaLeuGlu 502
      |||  :|||  |||  :|||
Qy      1454  TGG-----CACGATCCCGCACCCGAAACCTACCGAACACCGGACCGCAAGCAAACTT 1507
      |||  :|||  |||  :|||
Db      503  TrpThrGlyThrGluGlyGlnThrAlaThrTyrtHisValAspProHisAlaGlyLeu 522
      |||  :|||  |||  :|||
Qy      1508  TGGAAACGATCGCGCGCAAAATCTATCTGCTGCGCCATCGAAGGTTTATTA 1561
      |||  :|||  |||  :|||
Db      523  LeuArgAsnValMetThrGlyIlePheThrLeuLeuProValAspAspGlnLeu 540
      |||  :|||  |||  :|||
RESULT 40
Q88BC2_PSES
ID Q88BC2_PSES PRELIMINARY; PRT; 542 AA.
AC Q88BC2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=PSPT00095;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Winn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
DR EMBL; AE016853; AA053649.1; -; Genomic_DNA.
DR TIGR; PSPT00095; -.
DR GO; GO:0003824; F:metabolic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
SQ SEQUENCE 542 AA; 60833 MW; 8B96B8F00C45F7B8 CRC64;

```

```
Percent Similarity: 52.6% Conservative: 86
Best Local Similarity: 36.8% Mismatches: 187
Query Match: 27.0% Indels: 71
DB: 2 Gaps: 17

US-10-665-990A-13 (1-1561) x Q88BC2_PESM (1-542)

QY 68 TGCCTCTTCTCTGTTTCATGTTCTTCATGTTGCCCCCTGGAAGAACGACGGAAGC 127
Db 26 CysLeuLeuValAlaValCysValSer-GlyCysAlaHis-GlnGlnIleAlaSerGluP 45
QY 128 GGTCAATTTCAATCTTCCAAACCTGCTCTCTCT- - - - -CGACAACATCTCTG 172
Db 45 roser- - - - -GlnAlaMetProProSerAspSerAlaPheGlyArgSer-Ile 60
QY 173 CAAATCCGGGCACACCCCTCATAAACGGGTATCCGACATCTACTCTCTCGACGACCCC 232
Db 61 GlnAlaMetAlaMetProHis- - -GluGlyArgSerGlyPheArgLeuLeuProAspSer 79
QY 233 CACGAAGCCCTTGCCTGGCGCGCCCTTATCGAATCTGCGGAACACAGCCTCGATTG 292
Db 80 SerAspAlaPheLysAlaArgAlaGluLeuIleArgAsnAlaLysThrSerLeuAspLeu 99
QY 293 CAATACTACATTTGCGGAACGACATTTCCGCGAGCTGCTTCAACCTCATGTACCTT 352
Db 100 GlnTyrTyrIleValHisAspGlyLeuSerThrArgAlaLeuIleAspGluLeuLys 119
QY 353 GCGCGAGAACGCGCGTGCCTACGCTGCTGTTGGACGACAAACACACGCGCGGTG 412
Db 120 AlaAlaAspArgGlyValArgValArgIleLeuLeuAspAspThrThrSerAspGlyLeu 139
QY 413 GACGATCTCTGCTGCGCTCGACAGCCATCCCAATATCGAAGTCGGCTGTTCAACCCC 472
Db 140 AspGlnAlaIleAlaThrLeuAlaAlaHisProAsnIleGlnIleArgLeuPheAsnPro 159
QY 473 TTCGCTCCTACGAAA- - - - -TGGCGCGACTCGGCTACTCCACCGACTTCCC 520
Db 160 GlnAsnLeuGlyArgGluThrGlyValThrArgSerLeuGlyArgLeuMetAsnLeuSer 179
QY 521 CGCCTCAACCGCGCATGCACAAACATCTTTACCGCGCAACACCGCGCCACCATCTC 580
Db 180 ArgGlnHisArgArgMetHisAsnLysLeuTyrLeuAlaAspSerSerAlaAlaIleVal 199
QY 581 GCGGACGCAATATCGGCGACGAATCTTCAAGTCGGTGAGGACACCGTTTTCGCCGAC 640
Db 200 GlyGlyArgAsnLeuGlyAspGluTyrPheAspAlaGluProAsnLeuAsnPheThrAsp 219
QY 641 CTGGACATCTCTCGCCACCGGACGCTGCTGCGCGAAGTATCGCAGACTTCGACCGCTAC 700
Db 220 IleAspMetLeuSerValGlyProValAlaGluGlnLeuGlyHisSerPheAsnGlnTyr 239
QY 701 TGGGCAAGCCATTCGCCCCACAAACCCGCGCATCTCCGCGCGGCAACATCGGCAAG 760
Db 240 TrpAsnSer- - - - -ThrLeuSerLys 246
QY 761 GGTCTTCAAGCACTCGGATAC- - - - -AACGACGAAACATCCAGACACGCGCTCTG 811
Db 247 ProIleAspGlnPheMetTyrPheLeuProAspGlyGlnAspLeuAlaGluAlaArgGln 266
QY 812 CGCTACCGGAAACCGTCAACAGTCC- - - - -CCCTCTTACCAAAAAATA 856
Db 267 ArgLeuAspAspSerLeuGluGlnAlaHisGlnGlnHisLysAlaLeuTyrGluArgLeu 286
QY 857 CAGACGGGACGCATCCAC- - - - -TGGCAGAGC- - - - -GTCCAA 889
Db 287 MetAlaTyrLysThrGlnProArgMetLysThrTrpLeuAsnGluLeuValTrpAlaHis 306
QY 890 ACCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGG 949
Db 307 AsnGlnAlaLeuTrpAspAlaProThrLysValLeuAlaGlnGlyGluProAspProHis 326
QY 950 - - - - -ATTGCGGAGGCTGCAAGACGCGCTCAAAACAGCCCGCAAAAGCGCTCTATCTG 1003
```

Search completed: May 2, 2006, 05:24:25
Job time : 460 secs


```
QY 122 GAAAGCCGTCATTTCAATCTCCAAACCTGTCCTCTCGGACAAACATCTCGAAATCCGG 181
Db 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTCATAAACAGGGCTATCCGACATCTACCTGCTCGAGACCCCGACGAGCC 241
Db 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
QY 242 CTTCGGCCGCGCGCCCTTATCGAATCTGCGGAACACAGCCTCGATTTGCAATCTAC 301
Db 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCGAACAGCATTTCCGGCAGGCTGCTGTTCAACCTCATGTACTTGCACGAAA 361
Db 106 IleTyrAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
QY 362 CCGCGCGTGGCGTACGCTGCTGTTGGACGACACACACGCGCGGCTTGGACGATCTC 421
Db 126 ArgGlyValArgValArgLeuLeuLeuAspAspAsnAsnThrArgGlyLeuAspAspLeu 145
QY 422 CTGCTCGCCCTCGACAGGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCTA 481
Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATCGCGCACTCGGCTACCTGACCGGACTTCCCGCGCTCAACCGCGCATGCAAC 541
Db 166 ArgLysTyrArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgMethHis 185
QY 542 ACAAACTCTTTACCGCGACACCGCGCCACCATCTACTCGGCGGACGAATATCGCGCAC 601
Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAAGTCGGTGAGACACCGTTTTCGCGGACCTCGGACATCTCGCCACCGGC 661
Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTCGTCGGGAAGTATCGCACGACTTCGACCGCTACTCGGCAAGCGCATTCGCGCCAC 721
Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCACGCGCATCATCGCAGCGGCAACATCGCAAGGTCCTCAAGCACTCGGATAC 781
Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGACGAAACATCCAGACACCGCTCTCTCGCTACCGCGAACCGTTCGAACAGTCGCC 841
Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCTTACCAAAAATACAGACGGGACGCATCGACTGGCGAGCGTCCAAACCGCGCTGATC 901
Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATTCGCGGAGG 961
Db 306 SerAspAspProAlaLysGlyLeuAspArgAspArgGlySerProIleAlaGlyArg 325
QY 962 CTGCAAGACGCGCTCAAAACAGCCCGGAAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC 1021
Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CTTCAAAAATCGGGACAGACGCACTGGGCAAAACCTGGTGCAGGACGGCATAGACGTTACC 1081
Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCCTGACCACTCGCTACAGCGGACCGAGCTTGGCGCGTCCATTTCGGGCTACGTCAAA 1141
Db 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAAAACCGCTGCTCAAAGCCGCGCATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
```

```
QY 1202 GTCCCCGCCACAAAAGACAAAGGCGCTGACCGGAGCTCCGTTAACGAGCTCGATGCCAAA 1261
Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTTCATTGTGGACGGCAAAACGCACTCTTCATCGGCTCATTCACACTCGACCCCGTTCC 1321
Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCACGGCTCAATACCGAAATGGCGTCTGTCATCGAAAGCCCCCAAAATCGCAGAACAGATG 1381
Db 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGGATACCAACACCCGAAATACGCTTACCGGTTACCTCGACAAAACAC 1441
Db 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
QY 1442 AACCGCTGCAATGGCAGCATCCGCGCACCCGCAAAACCTACCCGAAACGAAACCCGAAGCC 1501
Db 486 AsnArgLeuGlnTyrHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTTGGAAACGATCGCGCAAAAATCTATCCCTGCTGCCCATCGAAGGTTTATTA 1561
Db 506 LysLeuTyrLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525
RESULT 2
B81083
cardiolipin synthetase family protein NMB1434 [imported] - Neisseria meningitidis (strai
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81083
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;
Hickey, E.K.; Haft, D.H.; Salzberg, S.B.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masigiani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: B81083
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-508 <TET>
A:Cross-references: UNIPROT:Q9JYU0; UNIPARC:UPI00000C46E9; GB:AE002494; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1434
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 1,93e-143 Length: 508
Score: 2573.00 Matches: 497
Percent Similarity: 98.8% Conservative: 5
Best Local Similarity: 97.8% Mismatches: 6
Query Match: 90.2% Indels: 0
DB: 2 Gaps: 0
US-10-665-990A-13 (1-1561) x B81083 (1-508)
QY 38 ATGAAAAACAGCAGCGCTCATTTCCCTTTATGCGCTCTCTCTGTTCTCATGTTCTTCATCG 97
Db 1 MetLysThrArgSerLeuIleSerLeuLeuCysLeuLeuLeuCysSerCysSerSerTrp 20
QY 98 TTGCCCCCCTCGAGAACGCGAAGCGTCATTTCAATCTTCCAACTTCCAAACCTGTCCTC 157
Db 21 LeuProProLeuGluGluArgThrGluSerArgHisPheAsnThrSerLysProValArg 40
QY 158 CTGGACAAACATCCTCAAATCCGCGACACCCCTCATAAACACGCGCTATCCGACATCTAC 217
Db 41 LeuAspAsnIleLeuGlnIleArgHisThrProHisThrAsnGlyLeuSerAspIleTyr 60
QY 218 CTGCTCGACGACCCCGACGAGCCCTTTCGCGCGCGCGCCCTTATCGAATCTGCCGAA 277
Db 61 LeuLeuAsnAspProHisGluAlaPheAlaArgAlaLeuIleGluSerAlaGlu 80
```

```
QY 278 CACAGCCTCGATTGCAATCTACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTC 337
Db 81 HisSerLeuAspLeuGlnTyrTyrIleTrpArgAsnAspIleSerGlyArgLeuLeuPhe 100
QY 338 AACCTCATGTACCTTGGCGCAGACGGCGGTGGCGTACGCTGCTGTTGGACGACAA 397
Db 101 AsnLeuValTyrLeuAlaAlaGluArgGlyValArgValArgLeuLeuLeuAspAsn 120
QY 398 AACACGCGCGGTTGGACGATCTCCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGT 457
Db 121 AsnThrArgGlyLeuAspAspLeuLeuAlaLeuAspSerHisProAsnIleGluVal 140
QY 458 CGCTGTTCAACCCCTTCGTCTCTAGCAAAATGGCGGCACCTCGGTACTCGACGACTC 517
Db 141 ArgLeuPheAsnProPheValLeuArgIleYsrPargAlaLeuGlyTyrLeuThrAspPhe 160
QY 518 CCCCGCCTCAACCGCGCATCACACAAATCCTTTACCGCGGACAAACCGCGCACCAT 577
Db 161 ProArgLeuAsnArgMetHisAsnLysSerPheThrAlaAspAsnArgAlaThrIle 180
QY 578 CTCGCGGACGCAATATCGCGACGAATACCTTCAAAGTCGTTGAGGACACCGTTCGCC 637
Db 181 LeuGlyGlyArgAsnIleGlyAspGluTyrPheLysValGlyGluAspThrValPheAla 200
QY 638 GACCTGGACATCTCGCCACCGCAGCGCTCGCGGAAGTATCGCAGACTTCGACCGC 697
Db 201 AspLeuAspIleLeuAlaThrGlySerValValGlyGluValSerHisAspPheAspArg 220
QY 698 TACTGGCAAGCCATTCCGCCACACCGCCAGGCATCATCCGAGGGGCAACATCGGC 757
Db 221 TyrTrpAlaSerHisSerAlaHisAsnAlaThrArgIleIleArgSerGlyAspIleGly 240
QY 758 AAGGCTCTTCAAGCACTCGGATACAAAGCAACCAATCCACACACGCGCTCTCGGTAC 817
Db 241 LysGlyLeuGlnAlaLeuGlyTyrAsnAspGluThrSerArgHisAlaLeuLeuArgTyr 260
QY 818 CGCGAAACCGTGAACAGTCGCCCTCTACAAAATAATACAGCGGACGCGATCGACTGG 877
Db 261 ArgGluThrValGluGlnSerProLeuTyrGlnLysIleGlnThrGlyCysIleAspTrp 280
QY 878 CAGAGCTCCAAACCCGCTGATCAGGCACACCCCTCGCAAGAGCTCGACCGGACCGC 937
Db 281 GlnSerValArgThrArgLeuIleSerAspAspProAlaLysGlyLeuAspArgAspArg 300
QY 938 CGCAAAACCGCGATTGCGGAGGCTGCAAGACGCGTCAAACAGCCGCAAAAAGGTC 997
Db 301 ArgLysProProIleAlaGlyArgLeuGlnAspAlaLeuLysGlnProGluLysSerVal 320
QY 998 TATCTGTTTCAACCTATTTCGTCCTTACAAAATCCGGCACAGACGCACTGGCAAAACTG 1057
Db 321 TyrLeuValSerProTyrPheValProThrLysSerGlyThrAspAlaLeuAlaLysLeu 340
QY 1058 GTGCAGACGCGCATAGAGTTACCGTCTGACCAACTCGCTACAGGCGACGCGTTGCC 1117
Db 341 ValGlnAspGlyIleAspValThrValLeuThrAsnSerLeuGlnAlaThrAspValAla 360
QY 1118 GCCGTCCATTCCGGCTACGTCAAATACCGAAACCGGTCTCAAAGCGGCGCATCAAATC 1177
Db 361 AlaValHisSerGlyTyrValLysTyrArgLysProLeuLeuLysAlaGlyIleLysLeu 380
QY 1178 TAGAGCTGCNAACCAACCATCGCTCCCGCCCAAAAAGACAAAGCGCTGACCGGACG 1237
Db 381 TyrGluLeuGlnProAsnHisAlaValProAlaThrLysAspLysGlyLeuThrGlySer 400
QY 1238 TCGTAACCGCTGCTGATGCCAAACCTTCAATTGTCGACGGCAACCGCATCTTCATCGGC 1297
Db 401 SerValThrSerLeuHisAlaLysThrPheIleValAspGlyLysArgIlePheIleGly 420
QY 1298 TCATTCAACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGCGCTCGTCATCGAA 1357
Db 421 SerPheAsnLeuAspProArgSerAlaArgLeuAsnThrGluMetGlyValValIleGlu 440
QY 1358 AGCCCCAAAATCGCAGAACAGATGGAGCGCACCCCTCGCGCATACCAACCCGGAATACGCC 1417
```

```
Db 441 SerProLysIleAlaGluGlnMetGluArgThrLeuAlaAspThrThrProAlaTyrAla 460
QY 1418 TACCGCGTTACCTTCGACAAACACACCGCCTGCAATGGCAGCATCCCGCCACCGAAAA 1477
Db 461 TyrArgValThrLeuAspArgHisAsnArgLeuGlnTrpHisAspProAlaThrArgLys 480
QY 1478 ACCTACCGCGAACCAACCGGAGCCAACTTTTGGAAACGATCGCGCAAAAATCCTATCC 1537
Db 481 ThrTyrProLanGluProGluAlaLysLeuTrpLysArgIleAlaAlaLysIleLeuSer 500
QY 1538 CTGCTGCCCATCGAAGGTTTATTA 1561
Db 501 LeuLeuProIleGluGlyLeuLeu 508
RESULT 3
C64847
YmdC protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: C64847
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64847
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <BLAT>
A:Cross-references: UNIPROT:P75919; UNIPARC:UPI000013B993; GB:AE000026; GB:U00096; NID
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ymdC
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 3,32e-57 Length: 493
Score: 1102.50 Matches: 221
Percent Similarity: 65.5% Conservative: 83
Best Local Similarity: 47.6% Mismatches: 143
Query Match: 38.7% Indels: 17
DB: 2 Gaps: 5
US-10-665-990A-13 (1-1561) x C64847 (1-493)
QY 200 GGGCTATCCGACATCTACTGCTCGACGACCCCGACGAGCCCTTGGCGCGCGCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaAlaArgTyrArg 56
QY 260 CTTATCGAATCTCGCGAACACAGCCTCGATTTCGAATACTACATTTTGGCGCAACGACATT 319
Db 57 LeuAlaGluMetAlaGluHisThrLeuAspValGlnTyrTyrIleTrpGlnAspAspMet 76
QY 320 TCCGGCAGGCTGCTGTTTCAACCTCATGTCCTTGGCCGACAGCGCGCTGCGGTACGC 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTTGGACACAAACACGCGCGGTTTGGACGATCTCTGCTGCGCTCGACAGC 439
Db 97 LeuLeuLeuAspAsnAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCAAGTCGCGCTTCAACCCCTTCTGCTACGCAATGGCGCGCATC 499
Db 117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGTCTACTGACCGACTTCCCGCCCTCAACCGCGGATGACACAAATCTTTTACCGCC 559
Db 137 GlyTyrIleThrAspPheSerArgLeuAsnArgArgMetHisAsnLysSerPheThrVal 156
QY 560 GACAACCGCGCCACCATCTACTCGCGGACGCAATATTCGGCGACGAATATCTTCAAGTCGT 619
Db 157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
```

```
QY 620 GAGGACACCGTTTCCGCGACCTGACATCTCTCGCCACCGGACGCGTCTGTCGGCGAAGTA 679
Db 177 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCAGCACTTCGACCGTACTGCGCAAGCCATTCCGCCCAACACGCGCGCATATC 739
Db 197 AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnGlnValLeu 216
QY 740 CGCAGCGCAACATCGCGAAGGGT-----CTTCAAGCATCGGA 778
Db 217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGAGCAACATCCAGACACGCGCTCTCGCTACCGCGAACAACGTCGAAACAGTCG 838
Db 234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAATATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCGCGCTG 898
Db 250 ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCGACACCCCTGCAAAAGACTCGACCGACCGCGCAACCCCGGATTCGCCGG 958
Db 270 LeuSerAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGCGCTCAACAGCCCGGAAAGCGTCTATCTCGTTTACCCCTATTTC 1018
Db 290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe 309
QY 1019 GTCCCTACAAAATCCGGCACACGACCTCGGCAAACTGGTCAGGACGGCATAGCGTT 1078
Db 310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCCTGACCAACTCGCTACAGCGACCGACGCTGGCGCGTCCATTCGGCTACGTC 1138
Db 330 AlaIleLeuThrAsnSerLeuAlaLysAspValAlaValAlaIleAlaGlyTyrAla 349
QY 1139 AATACCGAAACCGCTGTCAAGCGCGCATCAAACTCTAGACGTCGCAACCAACCAT 1198
Db 350 ArgTrpArgLysLeuLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCGCTCCCGCCACACA---AAACACAAAGCGCTGACCGGAGCTCCGTAACACGCTGCAT 1255
Db 370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCACAAACCTTCAATGTGACGGCAACGCATCTTCATCGGCTCATCAACCTCGACCCC 1315
Db 390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 CGTTCGCGACGGCTCAATACCAATGGCGTGTCTCATCGAAGCCCAAAATCGCAGAA 1375
Db 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
QY 1376 CAGATGGAGCGCACCTCCCGCATACACACCGCAATACGCTACCGGCTTACCTCGAC 1435
Db 430 LeuIleAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuAsp 449
QY 1436 AAACACAAACCGCTGCAATGGACGATCCCGCCACCGGAAAA-----ACCTACCGCGAAC 1489
Db 450 ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluIleLeuLysLys 469
QY 1490 GAACCGGAAGCCAACTTGGAAACGCATCGCGCGGCAAAATCCTATCCCTGTCGCCCATC 1549
Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGGTTTATTATA 1561
Db 490 GluTrpLeuLeu 493
RESULT 4
H90806
probable synthase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C.Species: Escherichia coli
```

```
C.Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C.Accession: H90806
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A.Reference number: A99629; MUID:21156231; PMID:11258796
C.Accession: H90806
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-493 <HAY>
A.Cross-references: UNIPROT:Q8X9I7; UNIPARC:UPI00000D07FD; GB:BA000007; PIDN:BAB34847.1;
A.Experimental source: strain O157:H7, substrain RIMD 0509952
C.Genetics:
A.Gene: ECs1424
C.Superfamily: cardiolipin synthase
```

```
Alignment Scores:
Pred. No.: 4,35e-57 Length: 493
Score: 1100.50 Matches: 220
Percent Similarity: 65.5% Conservative: 84
Best Local Similarity: 47.4% Mismatches: 143
Query Match: 38.6% Indels: 17
DB: 2 Gaps: 5
```

US-10-665-990A-13 (1-1561) x H90806 (1-493)

```
QY 200 GGGGTATCCGATCTACTCTCTGACAGACCCCGCAAGAGCGCTTGGCGCGCGCCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaPheAlaAargTyrArg 56
QY 260 CTTATCGAATCTCGGAACACAGCTCGATTGCAATACTACATTGGCGCAACGACATT 319
Db 57 LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrTrpIleTrpGlnAspMet 76
QY 320 TCCGCGAGCTGCTGTTCACCTCATCTGTACCTTGGCGAGAACGGCGCTGCGCTAGCG 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTTGGACGACAAACACACGCGGGTTGACGATCTCTGCTCGCCCTCGACAGC 439
Db 97 LeuLeuLeuAspAspAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCATATCGAAGTCGCGCTGTCAACCCCTTCCTCTACGCAAAATGGCGCGCATC 499
Db 117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGCACAAACAAATCCTTTACCGCC 559
Db 137 GlyTyrIleThrAspPheSerArgLeuAsnArgArgMetHisAsnLysSerPheThrVal 156
QY 550 GACAAACCGCGCCACCATCTCGCGGACGCAATATCGCGGACGAATATCTCAAGTCGGT 619
Db 157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
QY 620 GAGGACACCGTTTTCGCGACCTGGACATCTCGCACCGCGGACGCGTCTCGCGCAAGTA 679
Db 177 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCAGCACTTCGACCGCTACTGGGCAAGCCATTCGCCCCACACGCGCGCATATC 739
Db 197 AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnGlnValLeu 216
QY 740 CGCAGCGCAACATCGCGAAGGGT-----CTTCAAGCATCGGA 778
Db 217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGAGCAACATCCAGACACGCGCTCTCGCTACCGCGAACAACGTCGAAACAGTCG 838
Db 234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAATATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCGCGCTG 898
Db 250 ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCGACACCCCTGCAAAAGACTCGACCGACCGCGCAACCCCGGATTCGCCGG 958
Db 270 LeuSerAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGCGCTCAACAGCCCGGAAAGCGTCTATCTCGTTTACCCCTATTTC 1018
Db 290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe 309
QY 1019 GTCCCTACAAAATCCGGCACACGACCTCGGCAAACTGGTCAGGACGGCATAGCGTT 1078
Db 310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCCTGACCAACTCGCTACAGCGACCGACGCTGGCGCGTCCATTCGGCTACGTC 1138
Db 330 AlaIleLeuThrAsnSerLeuAlaLysAspValAlaValAlaIleAlaGlyTyrAla 349
QY 1139 AATACCGAAACCGCTGTCAAGCGCGCATCAAACTCTAGACGTCGCAACCAACCAT 1198
Db 350 ArgTrpArgLysLeuLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCGCTCCCGCCACACA---AAACACAAAGCGCTGACCGGAGCTCCGTAACACGCTGCAT 1255
Db 370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCACAAACCTTCAATGTGACGGCAACGCATCTTCATCGGCTCATCAACCTCGACCCC 1315
Db 390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 CGTTCGCGACGGCTCAATACCAATGGCGTGTCTCATCGAAGCCCAAAATCGCAGAA 1375
Db 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
QY 1376 CAGATGGAGCGCACCTCCCGCATACACACCGCAATACGCTACCGGCTTACCTCGAC 1435
Db 430 LeuIleAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuAsp 449
QY 1436 AAACACAAACCGCTGCAATGGACGATCCCGCCACCGGAAAA-----ACCTACCGCGAAC 1489
Db 450 ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluIleLeuLysLys 469
QY 1490 GAACCGGAAGCCAACTTGGAAACGCATCGCGCGGCAAAATCCTATCCCTGTCGCCCATC 1549
Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGGTTTATTATA 1561
Db 490 GluTrpLeuLeu 493
```

Db	250	ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu	269
QY	899	ATCAGCGACACCCCTCGAAAGAGATCGACCGCGACCGCGCAAAACCCGCGATTCCCGGG	958
Db	270	LeuSerAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln	289
QY	959	AGGCTGCAAGCGCTCAACAGCGCGGAAAGCGTCTATCTGGTTTCAACCTATTTC	1018
Db	290	ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleIleSerSerTyrPhe	309
QY	1019	GTCCCTACAAATCCGCGCACAGCCACTGCGCAAACTGGTCGAGGACGGATAGACGTT	1078
Db	310	ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle	329
QY	1079	ACCGTCTGACCAACTCGTACAGCGCGACGCGTGGCGGCTCCATTCGCGGTACGTC	1138
Db	330	AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGlyTyrAla	349
QY	1139	AAATACCGAAACCGCTGCTCAAGCCCGCATCAAACTCTACGAGCTCAACCCCAACCAT	1198
Db	350	ArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg	369
QY	1199	GCGGTCCCGGCACA--AAAGACAAAGCGCTGACCGCGAGCTCCGTAAACGAGCTGCAT	1255
Db	370	GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis	389
QY	1256	GCCAAACCTTCATTGTGAGCGGCAACGCATCTTCATCGGCTCATTCACACCTCGACCC	1315
Db	390	AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro	409
QY	1316	CGTTCCGCGACGGCTCAATACCGAAATGGCGTGTGATCGAAAGCCCAAAATCCGCGAA	1375
Db	410	ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln	429
QY	1376	CAGATGGAGCCACCTCGCGATACACACCGCAATACGCTACCGGTTACCTCGAC	1435
Db	430	LeuIleAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuAsp	449
QY	1436	AAACACAAACCGCTGCAATGCGACGATCCCGCCACCCGAAAA-----ACCTACCCGAAC	1489
Db	450	ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluIleValLeuLysLys	469
QY	1490	GAACCGGAGCGCAAACTTGTGAACGCGATCCCGCGCAAAATCTATCCCTGCTCCCATC	1549
Db	470	GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal	489
QY	1550	GAAGGTTATTA 1561	
Db	490	GluTrpLeuLeu 493	
RESULT 5			
D85666			
Probable synthase ymdC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)			
C;Species: Escherichia coli			
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004			
C;Accession: D85666			
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew			
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouisis, K.; Apodaca,			
Nature 409, 529-533, 2001			
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.			
A;Reference number: A85480; MUID:21074935; PMID:11206551			
A;Accession: D85666			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-493 <STO>			
A;Cross-references: UNIPROT_Q8X9I7; UNIPARC:UPI0000165766; GB:A8005174; NID:gl2514578; E			
A;Experimental source: strain O157:H7, substrain EDL933			
C;Genetics:			
A;Gene: ymdC			
C;Superfamily: cardiolipin synthase			
Alignment Scores:			
Pred. No.: 1.12e-56 Length: 493			
Score: 1093.50 Matches: 219			
Percent Similarity: 65.3% Conservative: 84			
Best Local Similarity: 47.2% Mismatches: 144			
Query Match: 38.3% Indels: 17			
DB: 2 Gaps: 5			
US-10-665-990A-13 (1-1561) x D85666 (1-493)			
QY	200	GGGCTATCCGACATCTACCTCTCGACGACCCCGACAGCCCTTGCCTGCGCGCGCCGCC	259
Db	37	GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaPheAlaAlaArgTyrArg	56
QY	260	CTTATCGAATCTGCGCAACACAGCTCGATTTCGAATACTACTATTGGCGCAACACAT	319
Db	57	LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrIleTrpGlnAspMet	76
QY	320	TCCGGCAGGCTGTGTTCAACCTCATCTCCGCGAGAACCGCGGTGCGGTACGC	379
Db	77	SerGlyArgLeuLeuPheSerAlaLeuAlaLysArgGlyValArgValArg	96
QY	380	CTGCTGTGGACGACAAACACCGCGGGTGGACGATCTCTCTGCTCGCTCGACAC	439
Db	97	LeuLeuLeuAspAspAsnAsnThrProGlyLeuAspIleLeuArgLeuLeuAspSer	116
QY	440	CATCCCAATATCGAAGTGCCTCTTCAACCCCTTCGCTCTACGCAAAATGCGCGAC	499
Db	117	HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu	136
QY	500	GGCTACTGACGCTTCCCGCTCAACCGCGCATGCAACAACAATCTTTACCGCC	559
Db	137	GlyTyrIleThrAspPheSerArgLeuAsnArgArgMethHisAsnLysSerPheThrVal	156
QY	560	GACAAACCGCCACCATCTCGCGGACGCAATATCGCGACGCAATACTTCAAAGTCGT	619
Db	157	AspGlyValValThrLeuValGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly	176
QY	620	GAGGACACCGTTCGCGGACCTCGACATCTCTCCACCGCGACGCTCGCGCAAGTA	679
Db	177	GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal	196
QY	680	TCGACACGCTTTCGACCGCTACTGGCAAGCATTCGCCCCACACGCGCGCATCATC	739
Db	197	AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnValLeu	216
QY	740	CGCAGCGCAACATCGCAAGGCT-----CCTCAAGCACTCGGA 778	
Db	217	-----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp	233
QY	779	TACAAACGACGAAACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCAACAGTCG	838
Db	234	HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer	249
QY	839	CCCCTTACCAAAATATACAGCGGACGATCGACTGGCAGAGCGTCCAAACCGCGCTG	898
Db	250	ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu	269
QY	899	ATCAGCGACACCTCGCAAAAGACTCGACCGCGACCGCGCAACCGCGATTGCCGGG	958
Db	270	LeuSerAspProAlaLysGlyGlyGlyLysAlaLysArgHisSerLeuLeuProGln	289
QY	959	AGGCTGCAAGCGCTCAAAACAGCCCGAAAAAGCGTCTATCTGTTTCCACCTATTTC	1018
Db	290	ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleIleSerSerTyrPhe	309
QY	1019	GTCCCTACAAATCCGCGCACAGCCACTGCGCAAACTGTGTGAGGACGCGATAGAGTT	1078
Db	310	ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLys**ValLysIle	329
QY	1079	ACCGTCTGACCAACTCGCTACAGGCGACCGAGTTCGCCCGCTCCATTCGGGTAGTC	1138
Db	330	AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGlyTyrAla	349
QY	1139	AAATACCGAAACCGCTGCTCAAGCCCGCATCAAACTCTACGAGCTCAACCCCAAC	1198

conserved hypothetical protein STY1185 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004
C:Accession: AE0636
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Th., T.; Connor, P.; Cronin, A.; Davis, R.M.; Dowd, L.; White, N.; Farrar,
S.S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.A;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; UID:2153494; PMID:11677608
A:Accession: AE0636
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-495 <PAR>
A:Cross-references: UNIPARC:UIP0000059FC8; GB:AL513382; PIDN:CAD08272.1; PID:g16502319;
C:Genetics:
A:Gene: STY1185
C:Superfamily: cardiolipin synthase

Alignment Scores:

Pred.	No. :	Length:
Fold	Score	Matches:
Percent Similarity:	63.9%	Conservative: 77
Best Local Similarity:	47.4%	Mismatch: 147
Query Match:	38.2%	Indels: 21
DB:	2	Gaps: 5

US-10-665-990A-13 (1-1561) x AE0636 (1-495)

Qy 200 GGGCTATCCGACATCTACTCTCGAGCACCCCCAACGAAGCCCTTGC CGCGCCGCCGCC 259

D b 39 GlyGlucysGlyIleLeuAlaLeuAsnSerLeuaspAlaPheAlaArgTyrrArg 58

Qy 260 CTTATCGAATTCGCGCAACACAGCTGTATTGTAATACTACATTTGGCGCAACGACATT 319

D b 59 LeuthrGluMetAlaAlaArgThrLeuAspValGlnTrpLysGlyIleLeuProVal 78

Ov 320 TC CGCGAGGCTGCTTTCAACCCTCATGTACCTTCGCGCAGAACCGCGCTGCGGTACGC 379

Db 430 AlaThrLeuIleHisLysArgPheThrGlnSerGlnArgAspAlaAlaTrpGlnLeuArg 449
 Qy 1430 CTCGACAAACAAACCGCTGCAATAGGCACGAT-----CCGCCACCCGAAACCTAC 1483
 Db 450 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGluGluLysValLeu 469
 Qy 1484 CCGAACGACCGGAGCGAACTTGGAAAGCGATCGCGCGCAAAATCTCTATCCCTGCTG 1543
 Db 470 LysLysGluProAlaThrArgPheTrpGlnArgValLeuValArgLeuAlaAilaLeu 489
 Qy 1544 CCCATCGAAGGTTTATTA 1561
 Db 490 ProValGluTrpLeuLeu 495
 RESULT 7
 G97556
 hypothetical 55.9K protein in csge-mdog intergenic region [imported] - Agrobacterium tum
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: G97556
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: G97556
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-518 <KUR>
 A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI00000D1C0B; GB:AE007869; PIDN:AAK87408.1;
 C:Genetics:
 A:Map position: circular chromosome
 C:Superfamily: cardiolipin synthase
 Alignment Scores:
 Pred. No.: 7,12e-40 Length: 518
 Score: 807.00 Matches: 204
 Percent Similarity: 54.6% Conservative: 85
 Best Local Similarity: 38.6% Mismatches: 196
 Query Match: 28.3% Indels: 44
 DB: 2 Gaps: 16
 US-10-665-990A-13 (1-1561) x G97556 (1-518)
 Qy 53 CTCATTTCCTTTATGCTCTCTCTCTGTTTCATGTTCTTCATGGTGGCCCACTGGAA 112
 Db 8 IleIleLeuLeuMetLeuIleGlyProSerLeuPheValIleGlyLysGlnArg 27
 Qy 113 GAACGAGCGGAAAGCGCTATTTCAATACTTCCAAACCTGTCTCTCGGACAAACATCCTG 172
 Db 28 GluLysAlaIleProLysArgProSerThrAlaLeuProValThrGluAspGluThrAla 47
 Qy 173 CAATCCGGACACCCCTCAT----AACAGGGGTATCCGAC-----ATCTACCTG 220
 Db 48 LeuAspArgHisTrpGlnSerIleArgAsnGlyTrpAsnGluLysAsnAlaLeuCysLeu 67
 Qy 221 CTGACGACCCCGAGAGCCTTGGCGCGCGCCGCTTATCGAATCTGCGGACAC 280
 Db 68 LeuHisSerAsnLeuAspAlaPheAlaValArgValAlaAlaAlaGlyArg 87
 Qy 281 AGCTCGATTTGCAATACTACTATTGGCGCAACGACATTTCCGCGAGGCTGTTTCAAC 340
 Db 88 SerLeuAspLeuMetTyTrpMetTrpAsnAlaAspLeuThrGlyArgLeuMetMetArg 107
 Qy 341 CTCATGTACCTTGGCGAGAACCGCGGTGCGGTACGCTGTGTTGGACGACAAAC 400
 Db 108 GluValIleAlaAlaAspArgGlyValArgValArgLeuLeuLeuAspLeuGly 127
 Qy 401 ACGCGGGTGGACGATCTCTGCTCGCCTCGACGACCATCCCATATCGAAGTGGC 460
 Db 128 ValSerMetSerAspArgIlePheHisAlaIleAspSerHisProAsnIleGluLeuArg 147

Qy 461 CTGTTCAACCCCTTCGTCCTACGCCAAA-----TGGCGCGCACTCGGTACTCGACC 511
 Db 148 LeuPheAsnProThrArgAlaArgGluAsnIleLeuHisArgSerLeuGluLeuValLeu 167
 Qy 512 GACTTCCCGCGCTCAACCGCGCATGCACAAATAATCTTTTCCGCGCAACACCGGCC 571
 Db 168 ArgPheArgSerValAsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgAla 187
 Qy 572 ACCATCTCGCGGACCAATATCGCGCGCAATACCTTCAAGTCGTCGAGCACCCGTT 631
 Db 188 ValIleValGlyArgAsnIleGlyAspAlaTrpPheAspAlaAlaGluArgAlaAsn 207
 Qy 632 TTCGCCGACCTGGACATCTCGCCACCGCGAGCGCTCGTCGCGCAAGATATCGCAGACTTC 691
 Db 208 PheHisAspPheAspIleLeuGlyPheGlyGlyIleValAlaAspAlaThrGluIlePhe 227
 Qy 692 GACCGTACTGGGAACCCATTCGCGCCCAAC-----GCCACG 730
 Db 228 AspAspTyTrp-----AsnSerAlaValSerValProValArgSerLeuLeuAlaArg 245
 Qy 731 CGCATCATCGCAGCGCAACATCGGCAAGGTTCTTCAAGCACTCGGATACACGACGAA 790
 Db 246 ArgProAsnLysLeuAlaLysLeuArgGluLeuAspAlaLeuProGlnSerGlu--- 264
 Qy 791 ACATCCAGACACGCGCTCTCGCTACCGCAACCGTCGAAACAGTCGCGCCTCTACCAA 850
 Db 265 -----AlaAlaArgProTyLeuGluArgValGluSerGln-----TyGly 278
 Qy 851 AAA-----ATACAGAGGAGCGCATCGACTGGCAGCGCTTCAAAACCCGC-----CTG 898
 Db 279 ArgAspHisPheLeuMetSerAspArgLeuHisTrp-----ValAspThrAlaAspVal 296
 Qy 899 ATCAGGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCCG-----ATT 952
 Db 297 LeuAlaAspProProGluLysAlaAlaGlyLys---ArgArgLysGlyHisAsnPheLeu 315
 Qy 953 GCCGGGAGGTGCAAGACGCGCTCAAAACAGCCCGCAAAACCGTCTATCTGTTTCAACC 1012
 Db 316 MetGluSerLeuLeuProLeuMetGlnAlaAlaGlyGluSerLeuHisIleThrSerPro 335
 Qy 1013 TATTTCGCTCATAAAATCCGGACAGACGCACTGGCAAACTGTGTGAGGAGCGGATA 1072
 Db 336 TyrPheIleProGlyLysGlnGlyValGluIlePheLeuAspLeuAlaGluArgGlyVal 355
 Qy 1073 GAGGTACCGTCTGACCACTCGTACAGCGACCGACGCTGCGCGCTTCCATTCGCGC 1132
 Db 356 SerLeuAlaIleLeuThrAsnSerLeuAlaAlaThrAspValAlaAlaValHisAlaGly 375
 Qy 1133 TAGCTCAAAATACCGAAACCGCTGCTCAAAAGCGGCATCAAACTCTACGAGCTGCAACCC 1192
 Db 376 TyrAlaArgTyArgLysProLeuLeuSerGlyGlyValArgLeuHisGluLeuArg--- 394
 Qy 1193 AACCATGCGCTCCCGCCCAAAAGAACAAAGGC-----CTGACCGGACGCTCGTA 1243
 Db 395 -----SerGlnAlaAspGlnGlySerPheThrLeuArgGlySerGlyGln 409
 Qy 1244 ACCAGCTCGATCCCAAAACCTTCATGTGGCGGCAACGCATCTTCATCGGCTCATTC 1303
 Db 410 AlaSerLeuHisThrLysAlaPheThrArgAspGlyGluThrGlyTyIleGlySerLeu 429
 Qy 1304 AACCTCGACCCCGTTCCGCGACGCTCAATACCGAAATGGCGTCGTCATCGAAAGCCCC 1363
 Db 430 AsnPheAspProArgSerAlaSerLeuAsnThrGluMetGlyValValPheAsnSerAla 449
 Qy 1364 AAAATCGCAACAGATGGAGCGCACCTTCGCGATACCAACCCGAA---TAGCGCTAC 1420
 Db 450 ProLeuValAlaAlaArgMetAspGluIlePheAlaGluIleArgArgThrMetSerPhe 469
 Qy 1421 CGGTTACCTCGACAAACACACACCGCTGCATGG-----CACGATCCCGCCACCCGA 1474
 Db 470 GluLeuAspIleAspSerAlaAsnArgIleValTrpMetThrGluGluArgGlyGlnPro 489
 Qy 1475 AAAACCTACCCGAAACCGAACCGCAAACTTTGGAAACGCATCGCGCGCAAAATCCTA 1534

```

Db      490  LysileTyrArgArgGluProAspAlaAlaileSerArgGilleAlaGlyileMet 509
Qy      1535  TCCCTGCTGCCCATCAAGGTTTATTA 1561
Db      510  ArgValLeuProLeuGluSerGlnLeu 518

RESULT 8
AB2777
phospholipase D family protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2777
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI00000D1C0B; GB:AE008688; PIDN:AAL42632.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu1630
A:Map position: circular chromosome
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 7,12e-40 Length: 518
Score: 807.00 Matches: 204
Percent Similarity: 54.6% Conservative: 85
Best Local Similarity: 38.6% Mismatches: 196
Query Match: 28.3% Indels: 44
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x AB2777 (1-518)

Qy      53  CTCATTCCCTTTATGCTCTCTCTGTTTCATGTTCTTCATGTTGCCCCCATCGGAA 112
Db      8  lleilleLeuLeuMetLeuileGlyProSerLeuPheValileGlyLysGlnArg 27
Qy      113  GAACGAGCGAAAGCGTCATTTCATTAATCTTCAAACTCTCTCTCTCGACAACTCCTG 172
Db      28  GluLysAlaileProLysArgProSerThrAlaLeuProValThrGluAspGluThrAla 47
Qy      173  CAAATCCCGCACACCCCTCAT---AACACGGGCTATCCGAC-----ATCTACCTG 220
Db      48  LeuAspArgHisTrpGlnSerileArgAsnGlyTrpAsnGluLysAsnAlaLeuCysLeu 67
Qy      221  CTCGAGGACCCCGACGAGACCTTCGCCCGCGCGCCCTTATCGAATCTCGCCGACAC 280
Db      68  LeuHisSerAsnLeuAspAlaPheAlaValArgValAlaAlaAlaArgAlaAlaGlyArg 87
Qy      281  AGCCTCGATTTCGAATACATTTGGCGCAACAGCATTTCCGCGAGGCTGCTGTTCAC 340
Db      88  SerLeuAspLeuMetTyrTrpMetTrpAsnAlaAspLeuThrGlyArgLeuMetMetArg 107
Qy      341  CTCATGTACTTTCGCGCAGAACCGCGGTGCGGTACGCTCTCTGTTGGACGACAAAC 400
Db      108  GluValileAlaAlaAspArgGlyValArgValArgLeuLeuLeuAspLeuGly 127
Qy      401  ACGCGGGTTCGAGCATCTCTGCTCGCCTCGACGCCATCCCAATATCGAAGTGC 460
Db      128  ValSerMetSerArgGilePheHisAlaileAspSerHisProAsnilleGluLeuArg 147
Qy      461  CTGTTCAACCCCTTCCTCTACGCAAA-----TGGCGGCGCACTCGGCTACCTGACC 511
Db      148  LeuPheAsnProThrArgAlaArgGluAsnilleLeuHisArgSerLeuValLeu 167
```

```

Qy      512  GACTTCCCTCCCTCAACCCGCGCATGCACAACTCTTTTACCGCGCAACACCGCGCC 571
Db      168  ArgPheArgSerValAsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgAla 187
Qy      572  ACCATCTCGCGGAGCGCAATATCGGACGAATACTTCAAAGTCGGTGAGGACACCGTT 631
Db      188  VallleValGlyGlyArgAsnilleGlyAspAlaTyrPheAspAlaAlaGluArgAlaAsn 207
Qy      632  TTCGCCGACCTTGGACATCTCTCCGCCACCGCAGCGTCGTCGCGGAGTATCGCAGACTTC 691
Db      208  PheHisaspPheAspilleLeuGlyPheGlyGlyileValAlaAspAlaThrGluilePhe 227
Qy      692  GACCGCTACTGGCAAGCCATTCGCCCCACAAAC-----GCCACG 730
Db      228  AspAspTyrTrp-----AsnSerAlaValSerValProValArgSerLeuLeuAlaArg 245
Qy      731  CGCATCATCCGCGGCGGCAACATCGGCAAGGTCCTTCAAGCACTCGGATACAACGACGAA 790
Db      246  ArgProAsnLysLeuAlaLysLeuArgArgGluLeuAspAlaLeuProGlnSerGlu--- 264
Qy      791  ACATCCAGACACGCGCTCTCTCGCTACCGCGAAACCGTCGAAACAGTCGCCCTCTTACCAA 850
Db      265  -----AlaAlaArgProTyrLeuGluArgValGluSerGln-----TyrGly 278
Qy      851  AAA-----ATACAGACGGGAGCATCGACTGGCAGAGCGTCCAAACCCGCG---CTG 898
Db      279  ArgAspHisPheLeuMetSerAspArgLeuHisTrp-----ValAspThrAlaAspVal 296
Qy      899  ATCAGCAGACACCCCTGCAAAAGGAGCTCGACCGCGACCGCGCAACACCGCGC-----ATT 952
Db      297  LeuAlaAspProProGluLysAlaAlaGlyLys---ArgArgLysGlyHisAsnPheLeu 315
Qy      953  GCCGGAGGCTGCAAGACGCGCTCAACACAGCCCGGAAAGAGCGTCTATCTGTTTCACCC 1012
Db      316  MetGluSerLeuLeuProLeuMetGlnAlaAlaGlyGluSerLeuHisileThrSerPro 335
Qy      1013  TATTTCGTCCTCAAAAATCCGGCACAGACGACCTGGCAAACTGGTCGACGAGCGGATA 1072
Db      336  TyrPheileProGlyLysGlnGlyValGluilePheLeuAspLeuAlaGluArgGlyVal 355
Qy      1073  GAGCTTACGCTTCGACCACTCGCTACAGCGCGACGCTTGGCGCGCTCCATTCGCGC 1132
Db      356  SerLeuAlaileLeuThrAsnSerLeuAlaAlaThrAspValAlaAlaValHisAlaGly 375
Qy      1133  TACGTCAATACCGAAACCGCTGCTCAAGCCGCGCATCAAACTCTACGAGCTGCAACCC 1192
Db      376  TyrAlaArgTyrArgLysProLeuLeuSerGlyGlyValArgLeuHisGluLeuArg--- 394
Qy      1193  AACCATGCGCTCCCGCCACAAAGACAAAGGC-----CTGACCGGAGCTCCGTA 1243
Db      395  -----SerGlnAlaAspGlnGlySerPheThrLeuArgGlySerGlyGln 409
Qy      1244  ACCGCTGTCATGCCAAACCTTCTATTGTGACGCGCAACGATCTTCTCGGCTCATTC 1303
Db      410  AlaSerLeuHisThrLysAlaPheThrArgaspGlyGluThrGlyTyrlleGlySerLeu 429
Qy      1304  AACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGCTCATCGAAAGCCCC 1363
Db      430  AsnPheAspProArgSerAlaSerLeuAsnThrGluMetGlyValValPheAsnSerAla 449
Qy      1364  AAAATCGCAGAACAGATGAGGCGACCTCCCGCATACCACACCGCAA---TACGCCATC 1420
Db      450  ProLeuValAlaArgMetAspGluilePheAlaGluGluileArgArgThrMetSerPhe 469
Qy      1421  CGCGTTACCTCGCAAAACACAAACCGCTCGCAATGG-----CACGATCCCGCCACCGA 1474
Db      470  GluLeuAspilleAspSerAlaAsnArgileValTrpMetThrGluGluArgGlyGlnPro 489
Qy      1475  AAAACCTTACCGAACGAAACCGGAGCCAAACTTTGGAAACGATCCGCGCAAAATCTTA 1534
Db      490  LysileTyrArgArgGluProAspAlaAlaileSerArgGilleAlaGlyileMet 509
```

177	Db	HisAsnLysLeuTrpLeuAlaAspGlyThrAlaAlaIleValGlyGlyArgAsnLeuGly	196
599	Qy	GACGAATACTTCAAAGTCGGTGAGACACACCGTTTTCGGCGACCTGCAGCATCTCTCGCCACC	658
197	Db	AspGluTrpPheAsnAlaLysProGluMetAsnPheThrAspLeuAspLeuAspLeuGlyVal	216
659	Qy	GGCAGCGTCGTCGGCGAAGTATCGCACGACTTCGACCGCTACTCGGCAAGCCATTCGCGC	718
217	Db	GlyProIleAlaAsnGlnLeuSerHisSerPheAspGlnTyrTrpAsnSerAlaIleSer	236
719	Qy	CACAACGGCCACG-----CGCATCATCTCCGACGCGGCAACATCGCAAGGGTCTT	766
237	Db	ArgProIleGluAspPheLeuTrpArgAlaProTyrProGlyGluLeu-----	252
767	Qy	CAAGCACTCGGATACAACGACGAACATCCAGACACGCGCTCTCGGTAC-----CGC	820
253	Db	-----GluSerAlaArgLysLeuGlnArgTrpLeuArgLys	265
821	Qy	GA AACCGTCGAACAGTCG-----CCCTCTACCAAAAATAACAGACGGGACGC	868
266	Db	GluSerValLysGluSerGlyTyrIleArgHisLeuPheAspArgGlyAspGlnProArg	285
869	Qy	ATC---GACTGGCAGAGC-----GTCCAAACCCGCGCTGATCAGCAGCACCCCT	913
286	Db	LeuGlyAsnTrpLeuGluAsnLeuThrTrpAlaArgAlaGluAlaIleTrpAspAlaPro	305
914	Qy	GCAAAAGGACTCGACCGCGCGCCGCAACCGCCG-----ATTGCCGGAGGTGCCAA	967
306	Db	LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuLeuSerProHisLeuAla	325
968	Qy	GACGCGCTCAACACGCCGAAAAAGCGCTCTATCTGGTTTCACCCCTATTTCCTCCCTACA	102
326	Db	GlyLeuPheLysGlyValGlnLysGluLeuIleLeuValSerAlaTyrPheValProAla	345
1028	Qy	AAATCCGGCA CAGACGCACTGGCAAAACTGGTGCAGGACGGCATAGAGCTTACCGTCCTG	108
346	Db	LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu	365
1088	Qy	ACCAACTCGCTACAGGCGACCGAGTTCGCCCGCTCCATTCCGGCTAGCTCAAAATACCGA	114
366	Db	ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg	385
1148	Qy	AAACCGTGTCTCAAAGCGCGCATCAAACTCTACGAGCTGCAACCAACCATCGCTCCCC	120
386	Db	MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro	405
1208	Qy	GCCACA---AAAGACAAAGGCGCTGACCGCGAGTCCGTAAACGAGCTGCATGCCAAAACC	126
406	Db	LeuSerGlyAlaProTyrArgLeuHisGlySerSerSerAlaSerLeuHisSerLysAla	425
1265	Qy	TTCAATTGGACGGCAACCGCATCTTCATCGGCTCATTCACCTCGACCCCGTCCGCA	132
426	Db	MetValPheAspArgLysValPheIleGlySerPheAsnPheAspProArgSerIle	445
1325	Qy	CGGCTCAATACCGAAATGGGCGTCGTATCAAGAGCCCAAAATCGCAAGACAGATGGAG	138
446	Db	LeuTrpAsnThrGluValGlyValIleValAspSerProLeuLeuAlaGluGlnArg	465
1385	Qy	CGC---ACCTCTGGCGATACACACCCGGAATACGCTTACCGGTTACCTCGACAAACAC	144
466	Db	GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer	485
1442	Qy	AAC-----CGCCTGCAATGCAAGAT-----CCGCGCACCCCGAAAAACCTACCCG	148
486	Db	GlySerArgProLysLysLeuValTrpIleAspGluArgAspGlyArgAlaGlnValLeuArg	505
1487	Qy	AACGAACCCGAAGCAAACTTTGAAACCGCATCGCCGCAAAAAATCCTATCCCTGCTGCC	154
506	Db	HisGluPro---GlySerLeuTrpArgArgLeuAsnAlaTrpValAlaGlyMetIleGly	524
1547	Qy	ATCGAAGGTTTATTA	1561

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: Xf1087
C:Superfamily: cardioliipin synthase

Alignment Scores:

Pred. No.: 2,76e-18 Length: 652
Score: 439.00 Matches: 161
Percent Similarity: 40.1% Conservative: 69
Best Local Similarity: 28.0% Mismatches: 200
Query Match: 15.4% Indels: 144
DB: 2 Gaps: 21

US-10-665-990A-13 (1-1561) x B82724 (1-652)

```
QY 125 AGCGTCAATTTCAATACTTCCAAACCTGTCTCTCGACACATCCTGCAAAATCGGCAC 184
||||| : : ||| |||||
Db 61 SerArgAlaLeuSerGluSerThrProVal----- 70

QY 185 ACCCTCATACACGGCTATCCGACATCTACCTGCTCGAGACCCCGACGAGCCCTT 244
||||| : : ||| |||||
Db 71 ThrPro-----LeuHisTyrValThrIleLeuAspLysGlyGluSerAlaLeu 86

QY 245 GCGCGCGCGCGCGCTTATCGAATCTGCCGAACAGCAGCTCGATTTGCAATCTACAT 304
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 ValAlaArgIleAsnLeuIleArgSerAlaThrArgSerIleAspMetGlnThrTyrIle 106

QY 305 TGGCGCAACGACATTTCCGGCAGGCTGTGTTCACCTCATGTACTTCCGCGAGAACGC 364
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 PheAspLysAspAspSerAlaArgLeuIleMetAspGluLeuLeuThrAlaAlaArg 126

QY 365 GCGCTGCGGTACGCTGCTGTGACGACACACACGCGCGGTGGCGATCTCTCTG 424
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 GlyValArgValArgLeuIleAspGlnLeuSerAlaIleSerAspLeuAsnIleLeu 146

QY 425 CTCGCGCTC---GACAGCATCCCAATATCGAAGTGGCGCTGTTCAACCTCTCGCTA 481
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 GlyAlaLeuAlaGlyAlaHisValAsnPheGlnLeuArgIleTyrAsnProIlePheGly 166

QY 482 CGCAATGGCGGCGACTCGGCTAC-----CTGACCGCACTTCCCG 520
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 Lys-----AlaLysLeuAsnTyrGlyAspTyrValAlaSerValLeuCysCysPheArg 184

QY 521 CGCCTCAACCGCGCATGCACACAAATCTTTACCGCGCAACCGCGCCACCATCTC 580
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 ArgPheAsnGlnArgMetHisAsnLysLeuLeuValIleAspGluMetIleGlyVal 204

QY 581 GCGGACGCAATATCGGCGCAATATCTTCAAAGTCGGTGGAGCACACGTTTTCGCGC 640
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 GlyGlyArgAsnTyrGlnAspAspTyrTrpAspLeuGluTyrAsnPheArgAsp 224

QY 641 CTGGACATCTCGCCACCGGAGCGTGTGCGGGAAGTATCCGACGACTTCGACCGCTAC 700
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 225 ArgAspValLeuValAlaGlyProAlaValLeuGlnMetAlaValAsnPheAspAlaPhe 244

QY 701 TGGGCAAGCCATCTCGCCCGCACACCGCGCATCTCGCGCGGCAACATCGGCAAG 760
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 TrpAla-----AlaGluArgSerValProValGluCysLeuArgAspValGlyArg 261

QY 761 -----GGTCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACGCG 805
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 MetValLeuGlnAspGlyValProMetLeu-----ProSerAlaValPheHisPro 278

QY 806 CTCCTGCGC---TACGGGAACCGTCGAA---CAGTCGCCCTCTACCAAAA----- 853
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 LeuGluArgValGlnArgValThrAlaGluAlaAsnAspProAspTyrValLysArgThr 298

QY 854 -----ATACAGCGGAGCGCATCGACTGGCAGCGCTCCAAACCGCGCTG 898
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 299 PheValAspProAlaLeuAlaValAsnLysValGlnTyr----- 311
QY 899 ATCAGCGACACCCCTGCAAAAGAGTCTGCACGCGACCGCGCAAAACCG---CCGATTGCC 955
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 ValAlaAspLeuProGlnLysHisArgHisGluHisAsnAlaAsnProValSerIleGly 331
QY 956 GGGAGGCTGCAAGACGGCTCAAAACAGCCCCGAAAGAGCTCTATCTGGTTTACCCTAT 1015
||||| : : ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 ProArgLeuAspSerLeuIleSerAsnAlaArgHisGluValIleLeuGlnThrProTyr 351
QY 1016 TTCTGTCCTCAAAATCCGGCACAGCGACTGCGAAACCTG-----GTGACGAGCGC 1069
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 LeuValLeuSerLysProAlaLeuAsnIlePheArgLeuAsnArgThrGlnAspLys 371
QY 1070 ATAGACGTTTACCGTCTGACCACTCGCTACAGCGACCGCTGTGCGCGCTCCATTCC 1129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 ProArgValValValThrAsnSerLeuAlaIleThrAspAsnProIleValTyrAla 391
QY 1130 GGCTACGTCAATAC---CGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTG 1186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 LeuPheTyrLysTyrLysArgArgAsnMetArgAspLeuGlyPheAspIleTyrGluTyr 411
QY 1187 CAACCC----- 1192
Db 412 LysProPheProGlnAsnProProIleAspLeuThrGlyValValProMetAspGlyTrp 431
QY 1192 ----- 1192
Db 432 AsnAsnAspSerLeuGlnThrArgGlnAspLeuLeuAlaAlaLysValSerAspAsp 451
QY 1193 -----AACCATGCGTCCCGCCACAAAGACGAGCGCTGACCGGAGCTCCGTA 1243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 SerAspAsnAlaArgGlnLeuProSerLeuGlnAsnLysGlyGlnValAspArgVal 471
QY 1243 ----- 1243
Db 472 LeuArgThrGluThrArgProPheLeuGlyIleHisThrValAsnLysProLeuPro 491
QY 1244 -----ACCAGCTGATGCAAAACCTTCATTGTGTGACGGCAAA 1282
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 ValThrArgProGlyAlaArgMetGlyLeuHisAlaLysSerLeuValValAspArgLys 511
QY 1283 CGCATCTTCATCGGCTCATTCACCTGACCCCGTTCGCGACGCTCAATCCGAAATG 1342
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 512 ValGlyValIleGlyThrHisAsnPheAspProArgGlyGluSerTyrAsnThrGluAla 531
QY 1343 GCGCTGCTATCGAAAGCCCAAAATCGCAACAGATGGAGCGCACCCCTC---GCCGAT 1399
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 ValValValIleGluAspProAlaPheAlaArgLeuLeuAlaSerSerIleGluGlyAsp 551
QY 1400 ACCACACCGCAATACGCTACCGGTTACCTCGAACAACACACCGCTGCAATGGCAC 1459
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 552 MetAspProGlyAsnAlaTrpValVal----- 560
QY 1460 GATCCCGCACCCGAAACCTACCCG-----AACGAACCCGAGCCAAA 1504
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 -----AlaProArgLysLysLeuProGlyLeuTyrLysLeuAsnTyrSerValGlyLys 578
QY 1505 CTTTGGAAACGATCGCGCAAAATCCTATCCCTGTGTCGCC 1546
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 LeuSerGlu-----AlaLeuProValLeuAspLeuTrpPro 590
```

RESULT 13

A43932

mucin 2 precursor, intestinal - human (fragments)

N:Alternate names: mucin SMUC-41

C:Species: Homo sapiens (man)

C>Date: 10-Mar-1993 #sequence revision 12-Apr-1996 #text change 31-Dec-2004

C:Accession: A43963; A45106; B45106; A43932; B33532; A61257; P03328; P03329

J:Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim, Y.S.

J: Biol. Chem. 269, 2440-2446, 1994

A:Title: Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the

A:Reference number: A49963; MUID:94132002; PMID:8300571

A;Accession: A49963
A;Molecule type: mRNA
A;Residues: 1-639 <GUI>
A;Cross-references: UNIPROT:Q02817; UNIPROT:Q14884; UNIPARC:UPI0000177AE0; GB:L21998
R;Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Roche, E.M.; Lagace, R.E.; Kim, Y.S.
J. Biol. Chem. 267, 21375-21383, 1992
A;Title: The human MUC2 intestinal mucin has cysteine-rich subdomains located both upstr
A;Reference number: A45106; MUID:93016075; PMID:1400449
A;Accession: A45106
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 628-1895 <GUG>
A;Cross-references: UNIPARC:UPI000016AB12; GB:M94131; NID:9186395; PIDN:AAA59163.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIP:116706)
A;Accession: B45106
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 2037-3020 <GUG>
A;Cross-references: UNIPARC:UPI000016AB13; GB:M94132; NID:9186397; PIDN:AAA59164.1; PID:
A;Experimental source: colon
A;Note: sequence extracted from NCBI backbone (NCBIP:116698)
R;Toribara, N.W.; Gum Jr., J.R.; Culhane, P.J.; Lagace, R.E.; Hicks, J.W.; Petersen, G.M.
J. Clin. Invest. 88, 1005-1013, 1991
A;Title: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorp
A;Reference number: A43932; MUID:91358717; PMID:1885763
A;Accession: A43932
A;Molecule type: DNA
A;Residues: 1343-1350, 'L', 1352-1411, 'S', 1413-1448, 'P', 1450-1503, 'T', 1505-1915 <TOR>
A;Cross-references: UNIPARC:UPI000016AD85; GB:M74027; NID:9188863; PIDN:AAA59875.1; PID:
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIP:55749)
R;Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampert, D.T.A.; Kim, Y.S.
J. Biol. Chem. 264, 6480-6487, 1989
A;Title: Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and eviden
A;Reference number: A33532; MUID:89197956; PMID:2703501
A;Accession: B33532
A;Molecule type: mRNA
A;Residues: 1916-2193 <GUA>
A;Cross-references: UNIPARC:UPI00000731C8; GB:M22405; NID:9188873; PIDN:AAA36334.1; PID:
A;Experimental source: intestine
R;Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.; Basbaum, C.B.
J. Clin. Invest. 87, 77-82, 1991
A;Title: Human bronchus and intestine express the same mucin gene.
A;Reference number: A61257; MUID:91086481; PMID:1985113
A;Accession: A61257
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 'T', 1925-1948, 'TTS', 1952-1954 <JAN>
A;Cross-references: UNIPARC:UPI0000177AE1
A;Experimental source: bronchus
R;Xu, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang, D.; Jones, C.; Forstner,
Biochem. Biophys. Res. Commun. 183, 821-828, 1992
A;Title: Human intestinal mucin-like protein (MLP) is homologous with rat MLP in the C-t
A;Reference number: FQ0328; MUID:92198477; PMID:1550588
A;Accession: FQ0328
A;Molecule type: mRNA
A;Residues: 2328-2468 <XUG>
A;Cross-references: UNIPARC:UPI0000177AE2; GB:M86523
A;Experimental source: small intestine
A;Accession: FQ0329
A;Molecule type: protein
A;Residues: 2328-2342, 'K', 2344-2354 <XUG1>
A;Cross-references: UNIPARC:UPI0000177AE3
C;Genetics:
A;Gene: GDB:MUC2
A;Cross-references: GDB:120203; OMIM:158370
A;Map position: lip15.5-1lip15.5
C;Superfamily: von Willebrand factor type C repeat homology
C;Keywords: glycoprotein; intestine; tandem repeat
F;2766-2834/Domain: von Willebrand factor type C repeat homology <VWC>
Alignment Scores:
Pred. No.: 3.17e-12 Length: 3020

Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 2 Gaps: 22
US-10-665-990A-13 (1-1561) x A43932 (1-3020)
QY 119 ACGGAAAGCCGTCATTTCAAT-----ACTTCCAAACCTGCTCTCTGGAGAAC 166
Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGTCAAATCCGGCAGCACCCCTCATAAACAACGGGTATCCGACATCTACTGCTGCAC 226
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
QY 227 GACCCCAAGAGCCCTTGGCGCGCGCCCTTATCGAATCTGCCGAACACAGCCCTC 286
Db 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTGCAATACATTTGGCGGCAACGACATTTCCGGCAGGCTGCTGTCAACCTCA-- 344
Db 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACTTGGCGGAGAACGGCGGTGGCGTACGCC----- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysGlyTyAspTyLysIleArgValAs 1387
QY 381 -TGTGTGTGG-----ACGACAACAAC 400
Db 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrThrProSe 1407
QY 401 ACGCGCGGGTGGACGATCTCTGCTCGCCTCGACAGCCATCCATATATGAAAGTGGCG 460
Db 1407 rProProProThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTTCGTCCTACGCAAAATGGCGGCACTCGGCTACCTGACGACTTCCCC 520
Db 1423 -----SerProProThrThrThrThrThrThrProProProThrThrThrProSerProPr 1441
QY 521 CGGCTCAACCGCGCATGCAACAATCTTTACCGCGGACAAACCGCGGACCATCATCTC 580
Db 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGCGAGCAATATCGCGAGCAATCTTCAAAGTCGTGGAGGACACCGTTTTCGCGAC 640
Db 1457 -----ProIleSerThrThrThrProProProProThrThrThrThrThrThrThr 1474
QY 641 CTGGACATCTCGCCACCGGACGTCGTGGGAGAGTATCGCAGCACTTCGAC----- 695
Db 1474 rThrThrProSerProProThrThrThrThrThrThrThrThrThrThrThrThrThrPr 1494
QY 696 -----GCTACTGGGCAAGCCATTCGCCCAACAACGCCACGGCGCATCATCCGAGCGGC 748
Db 1494 oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh 1514
QY 749 AACATCGGCAAGGGTCTTCAAGCACTCGGATACCAACGACGAAACATCCAGACGCGCTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGGTACCGGAAACCGTCGAAC-----AGTCGCCCCCTCTACCAAAAAATACAGACG 862


```
Db 404 uSerSerAlaProValThrSerThrThrGluSerSerAlaProValThrSe 424
QY 536 ATGCACAAAT-----CCTTACCGCGGACAAACCGCCACCATCTCGC 583
Db 424 rSerThrGluSerSerAlaProValThrSerThrThrGluSerSerAl 444
QY 584 GGACCAATATCGCGGACGAATCTCAAGTCG-----GTGAGGACACGGTTTCG-- 635
Db 444 aProValThrSerSerThrThrGluSerSerAlaProValProThrProSerSerSe 464
QY 635 ----- 635
Db 464 rThrThrGluSerSerAlaProValThrSerSerThrThrGluSerSerAlaPr 484
QY 636 ---CCGACCTGGACATCTCGCCACCGCGCGTCGTCGCGAGTATCGCAGCTTC 691
Db 484 oValProThrProSerSerThrThrGluSerSerAlaProValThrSerThr 504
QY 692 GACCGCTACTGGCGAAGCCATTCCGCCACACA----- 722
Db 504 rThrGluSerSerAlaProValProThrProSerSerThrThrGluSerSerSe 524
QY 723 -----ACGCCACCGCATCATCCGACGCGCAACATCGCGCAAGGCTTCCAA 769
Db 524 rAlaProAlaProThrProSerSerThrThrGluSerSerAlaProValThrSe 544
QY 770 GCAC-----TCGNATACAACGACGAAACA 793
Db 544 rSerThrThrGluSerSerAlaProValProThrProSerSerThrThrGluSe 564
QY 794 TCCAGACACGCGCTCTCGCTACCGCGCAACCGTCGAACAGTCGCCCTCTACCAAAA 853
Db 564 rSerThrProValThrSerSerThrThrGluSerSerSerAlaProValProThrP 584
QY 854 ATACAGCGGACGATCGATCTGGCAGAGCGTCCAAACCG--CCTGATCAGCGACAC 910
Db 584 roSerSerSerThrThrGluSerSerAlaProValProThrProSerSerThrT 604
QY 911 CCTGCAAAAGACTCGACCGCGACCGCGCAACCGCGATTCGCGGAGCTCGAAGC 970
Db 604 hrGluSerSerAlaProAlaProThrProSerSerSerThrThrGluSerSer 624
QY 971 CGCGCTAAACAGCCGCAAAAGCGCTCTATCTGGTTTCACCCCTATTTCGTCCTTACAAA 1030
Db 624 laProValThrSerSerThrThrGluSerSerSerAlaProValProThrProSerSer 643
QY 1031 TCCGGCACAGCGCACTGGCAAACTGGTGAGGACGATAGACGTTACGTCCTGACC 1090
Db 644 SerThrThrGluSerSerAlaProValProThrProSerSerThr-----Thr 661
QY 1091 AACTCGCTACAGGCGACCGATTCGCGCGC-----TTCAT 1126
Db 662 GluSerSerAlaProValProThrProSerSerSerThrThrGluSerSerAl 681
QY 1127 TCCGGCTACGTCAAATACCGAAACCGCTGCTCAAAAGCGCGCATCAAACTCAGC----- 1181
Db 681 aProValThrSerSerThrThrGluSerSerSerAlaProValThrSerSerThrGl 701
QY 1182 -AGTGCAACCCCAACCATCGCTCCCGCCCAAAAGACAAAGCCCTGACCGGAGCTCC 1240
Db 701 uSerSerAlaProValProThrProSerSerSerThrThrGluSerSerAlaPr 721
QY 1241 GTACACGCTGCATGCAAAACCTTCATTTGTGACGCGCAACGCATCTTCATCGGTCA 1300
Db 721 o-----ValProThrProSer-----SerSerThrThrGl 731
QY 1301 TTCAACCTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGCTCGTCATCGAAAGC 1360
Db 731 uSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerAlaPr 751
QY 1361 CCCAAAATCGCAGAACAGATGGACGCGACCTCGCCGATACCCACCCCGAATAGC----- 1415
```

```
751 oValThrSerSerThrThrGluSerSerSerAlaProValProThrProSerSerSerTh 771
1416 -----CCTACCGGGTTACCTCGACAAACCAACAC-----GCCTGCAA 1453
771 rThrGluSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSe 791
1454 TGGCACATCCCGCCACCGCAAAACCTACCGCAACGAAC-----CCGAA 1498
791 rAlaProValProThrProSerSerSerThrThrGluSerSerValAlaProValProTh 811
1499 GCCA-----AACTTTGAAACGATCCCGCAAAATCC 1532
811 rProSerSerSerSerAsnThrSerSerAlaProSerSerThr 826

RESULT 15
G70063
cardiolipin synthase homolog ywnE - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G70063
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni,
C.; Bron, S.; Brouillette, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C.
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallio,
iech, J.; Harwood, C.R.; Henaud, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinot,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue,
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, Y.; Pohl, T.M.; Portetelli,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G70063
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-482 <RUN>
A:Cross-references: UNIPROT:P71040; UNIPARC:UPI0000060B8F; GB:Z99122; GB:AL009126; NID
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywnE
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 5,78e-11 Length: 482
Score: 314.00 Matches: 126
Percent Similarity: 38.4% Conservative: 65
Best Local Similarity: 25.3% Mismatches: 189
Query Match: 11.0% Indels: 118
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x G70063 (1-482)
QY 113 GAACGACGAGAAAGCGTCATTTCAATCTTCCAAACCTCTCTCTGGGACCAACATCTG 172
Db 83 GluAspLeuGluThrLysGlnPheGlnPheAsnAsnArgAlaThrPheAspAsnLysAsp 102
QY 173 CAATCCGCGACACCCCTCATAAACACGCGCTA-----TCCGACATCTAC 217
Db 103 LeuIleTyMetLeuIleMetAsnAsnHisAlaValPheThrGluAspAsnSerValasp 122
QY 218 CTGCTCGACAGCCCGCCAGAGCCCTTGGCCGCGCGCCCTTATCGAATCTGCGCAA 277
Db 123 ValIleThrAspGlyArgAspLysPheGlnArgLeuLeuSerAspIleSerLysAlaLys 142
QY 278 CACAGCTCGATTTGCAATACTACATTTGGCGCAACGACATTTCCGCGCAGCTGCTTC 337
Db 143 AspHisIleHisLeuGlnIleTyIleTyLysGlyAspLeuGluLysLysLeuArg 162
QY 338 AACCTCATGTACCTTGGCGCAGAAACGCGGTGCGGTACGCTGCTGTTGGAGCACAAC 397
```

```
Db      163 AspAlaLeuIleGlnLysAlaLysGluGlyValGlnValArgValIleTyrAspGluLeu 182
      ::::  ::::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
Qy      398 AACAGCGCGGGTGTGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAGTG 457
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      183 GlySerArgThrLeuArgLysLysPhePheLysGluLeuArgGluAlaGlyHisVal 202
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      458 CGCTGTTCAACCCCTTCCTCTAGCAAAATGGCGGCACTCGGCTACCTGACCGACTTC 517
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      203 GluValPhePheProSerLysLeuArgProIleAsnLeu----- 215
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      518 CCCCGCTCAACCGCGCATGCACAAATCTTTACCGCGACAAACCGCGCCACATA 577
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      216 ----ArgLeuAsnTyrArgAsnHisArgLysLeuValIleAspGlyMetThrGlyTyr 234
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      578 CTCGGCGGACGAATATCGCGACCAATACTTC-----AAAGTCGGTGAGGAC 625
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      235 ValGlyGlyPheAsnValGlyAspGluTyrLeuGlyLeuAsnProLysPheGly----- 252
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      626 ACCGTTTCGGACCTGGACATCTCGCCACCGGAGCGTGTGGGGAAGTATCGCAC 685
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      253 ---TyrTrpArgAspThrHisIleArgLeuGlnGlyThrAlaValHisAlaIleGlnThr 271
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      686 GACTTCGACCGTACTGGGCAAGCCATTCGCCCAACACGCCCGCATCATCGCAGC 745
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      272 ArgPheIleLeuAspTrpAsnGlnAlaSerHisHisThrLeuThrTyrIleProAsn 291
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      746 -----GGCAACATCGGCAAGGTCCTTCAAGCACTCGGATAC 781
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      292 HisPheProAspTyrGlyProLysGlyAsnValGlyMetGlnIleValThrSerGlyPro 311
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      782 AACGACGAACATCCAGACACGCGTCTCGCTACCGCGCAACCGTCGAACAGTCGCC 841
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      312 AspSerGlu----- 314
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      842 CTCACCAAAATACAGACGGGACGATCGACTGCGAGAGCGTCCAAACCGCGCTGATC 901
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      315 ---TrpGluGlnIleLysAsnGlyTyrIle-----LysMetIle 326
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      902 AGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGGATTCGCGGAGG 961
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      327 SerAsnAla----- 329
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      962 CTGCAAGACGCGCTCAACAGCCCGGAAAGGTCATCTGTTTCAACCTATTTCGTC 1021
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      330 -----LysArgSerIleLeuIleGlnThrProTyrPheIle 341
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1022 CCTACAAAATCGGCACAGCAGCTGGCAAACTGGTGCAGGCGGCATAGACGTTACC 1081
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      342 ProAspAlaSerLeuLeuAspAlaLeuArgIleAlaCysLeuSerGlyIleAspValAsn 361
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1082 GTCCTGACCAACTCGCTACAGCGACCGAGTGTGGCGCGTCCATTTCGGGTACGTCAAA 1141
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      362 IleMetIleProAsn-----LysProAspHisAlaPheValTyrTrpAlaThrLeuSer 379
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1142 TACCGAAAACCGTGTCTAAAGCGCGCATCAACTCTACGAGTGCACCAACCATGCC 1201
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      380 TyrIleGlyAspLeuLeuLysAlaGlyAlaThrValTyr----- 392
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1202 GTCCCGCGCACAAAAGACAAAGCCGTGACCGCGAGCTCGTAAACAGCGCTGCATGCCAAA 1261
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      393 -----IleTyrAspAsnGlyPhe-----IleHisAlaLys 402
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1262 ACCTTCATTGTGAGCGGAAACGCACTCTTCATCGGCTCATCAACTCGACCCCGCTTC 1321
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      403 ThrIleValValAspAspGluIleAlaSerValGlyThrAlaAsnIleAspValArgSer 422
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1322 GCACGCTCAATACCAATGGCGTCTCATCGAAGCCCAAAATCCGACGAACAGATG 1381
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      423 PheArgLeuAsnPheGluValAsnAlaPheIleTyrAspIleThrIleAlaLysLysLeu 442
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1382 GAGCGCACCCCTCGCCGATACCAACCCCGAATACGCGCTACCGGCTACCCCTCGACAAACAC 1441
      ::::  |||  |||  |||  |||  |||  |||  |||
```

```
Db      443 ValSerThrPhe-----LysGlu 448
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1442 AACCGCTCGCAATGGCAGATCCGCCACCCGAAAA---ACCTACCCGAAACGCCGAA 1498
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      449 AspLeuLeu-----ValSerArgLysPheThrTyrGluGluTyrLeuGln 463
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      1499 GCCAACTTTGAAACGCATCGCCGCAAAAATCTATCCCTGCTG---CCCATC 1549
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      464 ArgProLeuTrpIleArgIleLysGluSerValSerArgLeuLeuSerProIle 481
      ::::  |||  |||  |||  |||  |||  |||  |||

RESULT 16
F90001
hypothetical protein SA1891 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90001
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-494 <KUR>
A:Cross-references: UNIPROT:Q99SG9; UNIPARC:UPI00000D778F; GB:BA000018; PID:g13701883; F:
A:Experimental source: strain N315
C:Genetics:
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 2,23e-10 Length: 494
Score: 304.00 Matches: 112
Percent Similarity: 39.0% Conservative: 68
Best Local Similarity: 24.3% Mismatches: 145
Query Match: 10.7% Indels: 136
Gaps: 17
DB: 2

US-10-665-990A-13 (1-1561) x F90001 (1-494)
Qy      134 TTCATACTCTCCAAACCTGCTCTCTGGACAAATCTCGCAATCTCGCAACACCCCTCAT 193
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      121 TyrAsnAlaAlaPheLeuThrThrAspAsnAspLeuLysIle---TyrThrAspGly 139
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      194 AACACGGGCTATCCGACATCTACGCTCGACGACCCCGACGAGCCCTTTCGCCCGCCG 253
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      140 GlnGluLysPheAspAsp-----LeuIleGlnAsp----- 149
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      284 GCGGCCCTTATCGAATCTGCCGACACAGCTCGATTGGCAATACTACATTTGGCGCAAC 313
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      150 -----IleArgAsnAlaThrAspTyrIleHisPheGlnTyrTyrIleGlnAsn 166
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      314 GACATTTCCGCGAGCTGCTGTCAACCTCATGTACCTTGGCGGAGACGCGCGCTGCGC 373
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      167 AspGluLeuGlyArgThrIleLeuAsnGluLeuGlyLysLysAlaGluGlnGlyValGlu 186
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      374 GTACGCCCTGCTGTTGGACGACAAACACGCGGGGTTGGACGATCTCTCTGCTCGCCCTC 433
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      187 ValLysIleLeuTyrAspAspMetGlySerArgGlyLeu----- 199
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      434 GACGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCCTCTAGCAAAATGGCGC 493
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      200 -----ArgLysLysGlyLeuArgProPheArgAsnLysGlyGlyHis 213
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      494 GCACCTCGGCTACTGACCGACTTCCCG-----CGCCTCAAC 529
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      214 AlaGluAlaPhe-----PheProSerLysLeuProLeuIleAsnLeuArgMetAsn 230
      ::::  |||  |||  |||  |||  |||  |||  |||
Qy      530 CGCCGCTGCAACAAATCTTTTACCGCCGACAAACCGCGCACCATCTCGCGGAGCGC 589
      ::::  |||  |||  |||  |||  |||  |||  |||
Db      231 AsnArgAsnHisArgLysIleValValIleAspGlyGlnIleGlyTyrValGlyPhe 250
      ::::  |||  |||  |||  |||  |||  |||  |||
```

```
QY 590 AATATCGCGCAGATACCTTCAAGTCGGTGAGGACACCGCTTTTCGCC-----GAC 640
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 251 AsnValGlyAspGluTyr-----LeuGlyLysSerLysPheGlyTyrTrpArgAsp 268
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 641 CTGACATCTCTCGCCACCGGACGCTCGTGGCGAAGTATCGCAGCATTCGACGGCTAC 700
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 269 ThrHisLeuArgGileValGlyAspAlaValAsnAlaLeuGlnLeuArgPheIleLeuAsp 288
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 701 TGGGCAACCATTCGCGCCCAACACGCGCGC-----733
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 289 TrpAsnSer-----GlnAlaThrArgAspHisIleSerTyrAspArgTyr 304
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 734 -----ATCATCCGAGCGCAACATCGCGAAGGTCTTCAAGCAGCTCGATACAAC 784
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 305 PheProAspValAsnSerGlyGlyThrIleGlyValGlnIleAlaSerSerGlyProAsp 324
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 785 GACGAAACATCGACACGCGCTCTGGCGTACCGGAAACCGTCGAAAGTCGCGCCCTC 844
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 325 GluGlu-----326
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 845 TACCAAAAATACAGACGGAGCGATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGC 904
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 327 TrpGluGlnIleLysTyrGlyTyrLeu-----LysMetIleSer 339
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 905 GACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGGAGGCTG 964
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 340 SerAla-----341
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 965 CAAGACGGCTCAACACGCCGAAAAAGGTCTATCTGTTTCAACCTATTTCGTCCT 1024
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 342 -----LysLysSerIleTyrIleGlnSerProTyrPheIlePro 354
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1025 ACAAATCCGGCACAGCGACTCGCAAAACTGTGTGAGGACGACATAGCTTACCGTC 1084
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 355 AspGlnAlaPheLeuAspSerIleLysIleAlaLeuGlyGlyValAspValAsnIle 374
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1085 CTGACCAACTCGCTACAGGCGACCGCGCTTCCCGCTAGCTCAAAATAC 1144
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 375 MetIleProAsn-----LysProAspHisProPheValPheTyrAlaThrLeuLysAsn 392
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1145 CGAAACCGCTGCTCAAGCGGATCAAACTTACGAGTGTGCAACCAACCATCGCGCTC 1204
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 393 AlaAlaSerLeuLeuAspAlaGlyValLysValPheHisTyr-----406
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1205 CCCGCCACAAAGACAAAGGCTGACCGGACGCTCCGTAAACCGCTGCGATGCCAAAC 1264
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 407 -----AspAsnGlyPhe-----LeuHisSerLysThr 415
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1265 TTCATTGTGGACGCAACGATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGCA 1324
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 416 LeuValIleAspAspGluIleAlaSerValGlyThrAlaAsnMetAspHisArgSerPhe 435
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1325 CGGCTCAATACCGAAATGGGGTGTGTCATGAAAGCCCAAAATCGCAGAACAGATGGAG 1384
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 436 ThrLeuAsnPheGluValAsnAlaPheIleTyrAspGlnIleAlaLysLysLeuLys 455
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1385 CGCACCTCGCGGAT-----ACCACACCGGAATACGCGCTAC 1420
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 456 GlnAlaPheIleAspLeuAlaValSerSerGluLeuThrLysAlaArgTyrAlaLys 475
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1421 CGC 1423
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 476 Arg 476
```

RESULT 17

cardiolipin synthase homolog lin0008 [imported] - *Listeria innocua* (strain Clip11262).
C;Species: *Listeria innocua*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: A11433
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madieno, E.; Maitournam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A;Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: A11433
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-504 <GLA>
A;Cross-references: UNIPROT:Q92FUS; UNIPARC:UPI00000CC0B3; GB:AL592022; PIDN:CAC95241.1
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0008
C;Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 3,84e-10 Length: 504
Score: 300.00 Matches: 118
Percent Similarity: 37.9% Conservative: 76
Best Local Similarity: 23.0% Mismatches: 210
Query Match: 10.5% Indels: 108
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x A11433 (1-504)

```
QY 26 ACTATATCCCGATGAAACACAGCGCTCATTTCCCTTTTATGCTCTCTCTGTTC 85
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 90 ThrAlaGlnValMetGluLysAlaLysLeuIleAsnAlaIleHisAlaIleProAsnAsn 109
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 86 TGTTCCTCATGTTGCCCCCACTGGAAGACGAGCGGAAACCGCTCATTTCAATATCTCC 145
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 110 ThrAsnGluLysLeuProArgLeuSerLysArgIle-----AlaHisLeuThrSerIle 127
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 AAACCTGTCTCTCGGACAAACATCTCTGCAAAATCGGCACACCCCTATAACAACGGCTA 205
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 128 GluProIle-----LysGly--- 132
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 206 TCCGACATCTACCTGTGACGACGCCCCACGAGCCCTTCCGCCCGCGCCCTTATC 265
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 133 AsnLysIleGluIleLeuThrAsnGlyGluGluThrPheProValLeuLeuAspAlaLeu 152
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 266 GAATCTCTGCCAACAACAGCTCGATTGTCATATCTACATTTGGCGCAACGACATTTCCGGC 325
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 153 ArgLysAlaGluAsnHisIleHisIleGlnTyrTyrIlePheLysThrAspAlaIleSer 172
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 326 AGGCTGTGTTCACCTCATGTACCTTGCAGAGAACGGCGGTGCGGTACGCTGCTG 385
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 173 ThrGluIleArgAspIleLeuValGluLysAlaLysSerGlyValGluValArgPheMet 192
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 386 TTGACGACACAACACGCGCGGTGGACGATCTCTGCTCGCTCGCTCGACGCCATCCC 445
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 193 PheAspGlyLeuGlySerSerLysLeuSerLysAlaPheLeuAlaProLeuLysGluAla 212
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 446 AATATCGAATCGCGCTGTTCACCCCTTCCGCTCTACGCAAAATCGCGCGCACCTCGCTAC 505
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 213 GlyValSerIleHisAlaPheAspPro---IleAlaSerProTyrIleVal-----228
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 506 CTGACCGACTTCCCGCGCTCAACCGCGCATCGACAACAATCTTTACCGCGCACAAC 565
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 229 -----ArgThrAlaAsnLeuArgAsnHisArgLysIleValIleAspGly 244
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 566 CGCGCCACCATACTCGCGGACGCAATATCGGCGACGAATAC-----607
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 245 GlnIleGlyPheThrGlyGlyLeuAsnIleGlyGluGluTyrArgSerAsnThrProAsp 264
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 608 TTCAAAAGTCGGTGAGGACACCGTTTTCGCGCATCTGGACATCTCGCCACCGCAGCGCTC 667
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 265 PheArgValTrpArgAspThr-----HisMetLysIleThrGlyGlnAla 279
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 668 GTCCGCGAAGTATCGCACGACTTC-----GACCGTACTGGCAAGCCATTCGCGCCAC 721
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 280 ValIleGluLeuGlnGluSerPheLeuAsnAspTyrIleTyrMetGluAsnGlnAlaGly 299
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



```
Db 242 ThrSerValGlnProSerSerSerSerSerAlaProThrThrThrSerAlaThrSerValGln 261
QY 825 CGTCTGAAAGTCGC-----CCCTTACCAAAATAACAGACGGACGCATCG 872
Db 262 ProSerSerSerSerThrProThrProThrThrThrSerValGlnProSerSer 281
QY 873 ACTGGCAGAGCGTCCAAACCCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCG 932
Db 282 -----SerSerAlaProThrThrSerAlaThrSerValGlnProSerSer--- 297
QY 933 ACCGCCGAACCCCGGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCGCAAAAAA 992
Db 298 -----SerSerProThrSer 303
QY 993 GCGTCTATCTGTTTCAACCTATTTCGTCCTACAAAATCGGCACACGACCTGGCAA 1052
Db 304 SerThrIleSerValGlnProSerSerSerSerSerProThrThrSer----- 320
QY 1053 AACTGGTCAGGACGGCATAGACGTTACCGTCTGACCAACTCGCTACAGGCGACCGACG 1112
Db 321 -----ThrThrSerValGlnProSer 327
QY 1113 TTGGCGCGCTCATTCGGCTACGTCAAATACC----- 1145
Db 328 SerSerGlySerAlaProThrThrSerAlaThrSerValGlnProSerSerSer 347
QY 1146 -----GAAACCCGCTGCTCAACCGCGCATCA 1172
Db 348 ProProIleSerSerThrIleSerValGlnProSerSerSerSerSerProThrThr 367
QY 1173 AACTCTAGAGCTGCAACCCCAACCATGCGCTCCCGCCACAAAAGACAAAGCGCTGACG 1232
Db 368 SerThrThrSerValGlnProSerSerSerGlySerAlaProThrThrSerAlaThrSer 387
QY 1233 GCAGCTCGTAACGAGCTGATGCCAAACCTTATTGTGGAGCGGCAAAAGCATCTCA 1292
Db 388 ValGlnPro-----SerSerSerSerValProThrThrThrSerAlaThrSer 403
QY 1293 -----TCGGCTCATTAACCTCGACCCCGCTCCGACGCTCAATACCGAATGGCG 1346
Db 404 ValArgSerSerSerSerSerSerProThrProThrThrThrSerValGlnProSer 423
QY 1347 TCGTCATCGAAAGCCCAAAATCG-----CAGAACAGATGAGCGCACCC 1391
Db 424 SerSerSerSerValProThrThrSerAlaThrSerValGlnThrSerSerSerSer 443
QY 1392 TCGCGGATACCA-----CACCGAATACGCTACCGGTTACCGTTCACCTCGACA 1436
Db 444 ThrProIleProSerThrThrSerValGlnProSerSerSerSerSerAlaProThrThr 463
QY 1437 AACACAACCGCTCG-----AATGGCAGATCCCGCACCCGCAAAACCTACCGAAG 1490
Db 464 SerAlaThrSerValGlnProSerSerSerSerSerProProIleSerSerThr---Ile 482
QY 1491 AACCCGAAGCAAACTTTGGAACGCGATCGCGCAAAATCC 1532
Db 483 SerValGlnProSerSerSerSerSerSerProThrThrSer 496
```

RESULT 19

T45462
membrane glycoprotein [imported] - equine herpesvirus 1
C:Species: equine herpesvirus 1
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45462
R:Kirisawa, R.; Kobayashi, T.; Kawakami, Y.; Iwai, H.
J. Equine Sci. 7, 79-87, 1996
A:Title: Nucleotide sequences of open reading frames 1, 24 and 71 of an attenuated equine herpesvirus 1
A:Reference number: 222973
A:Accession: T45462
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-866 <KIR>

A:Cross-references: UNIPROT:O39781; UNIPARC:UPI00000ECBA1; EMBL:D88733; PIDN:BAA20037.1
A:Experimental source: strain HH1
C:Genetics:
A:Note: ORF71
C:Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homology
F:558-866/Domain: equine herpesvirus 1 glycoprotein homology <EHG>

Alignment Scores:

Pred. No.:	5.6e-10	Length:	866
Score:	297.50	Matches:	122
Percent Similarity:	39.2%	Conservative:	72
Best Local Similarity:	24.6%	Mismatches:	222
Query Match:	10.4%	Indels:	79
DB:	2	Gaps:	12

US-10-665-990A-13 (1-1561) x T45462 (1-866)

```
QY 126 GCCGTCAATTTCAATACTTCCAAACCTGCTCTCTGGACAACATCTCTGCAATTCGGCACA 185
Db 19 AlaileGlySerThrThrThrThrThrGluThrThrThrSerSerSerSerThrSerGlySer 38
QY 186 CCCCTATAACAACGGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAGCCCTTG 245
Db 39 GlyGlnSerThrSerSerGlyThrThrAsnSerSerSerSerProThrThrSerPro--- 57
QY 246 CCGCCCGCGCGCCCTTATCGAATCTCGGAACACGACCTCGATTTCGAATACTACATTT 305
Db 58 -----ProThrThrSerSerSerProProThrThrThrHisThrSerSerProSer 74
QY 306 GCGCAACAGCATTTTCGCGGAGGCTGCTTAACTCATGCTACCTTCCGCGCAGAACGCG 365
Db 75 SerThrSerThrGlnSerSerSerSerThrAlaAlaThrSerSerSerSerAlaProSerThrAla 94
QY 366 GCGTCCGCTACGCTCTGTTGACAGACAACACGCGGGGTTGGACGATCTCTGCG 425
Db 95 -----SerSerThrSerIleProThrSerThrSerThr 106
QY 426 TCGCCCTCGACAGCATPCCCAATATCGAAGTGGCGCTTCAACCCCTTCGTCCTACGCA 485
Db 107 GluThrThrThrThrThrProThrAlaSerThrThrThrThrProThrThrThrThrAlaAla 126
QY 486 AATGGCGCGCACTCGGCTACCTGACGACTTCCCGCCCTCAACCGCGCATGCAACA 545
Db 127 -----ProThrThrAlaAlaThrThrThrAlaValThrThr 138
QY 546 AATCCTTTACCGCGCACACCGCGCCACCATATCTCGCGGACGCAATATCGGCGACGAAT 605
Db 139 -----AlaAlaSerThrSerAlaGluThrThrThrAlaThrAla 155
QY 606 ACTTCAAAGTCGGTAGGACACCGTTTTCGCCGACCTGGACATCCTCGCCACCGCGCAGCG 665
Db 156 ThrAlaThrSerThrProThrThrThrThrProThrSerThrThrThrThrThrThrAlaThr 175
QY 666 TCGTCCGCGAAGTATCGCAGCTTCGACGCTACTGGGCAAGCATTCCGCCCCACACAACG 725
Db 176 ThrThrValPro---ThrThrAlaSerThrThrThrThrThrThrThrThrThrThrThr 194
QY 726 CACGCGCATCATCCGACGCGGCAACATCGGAAGGCTCTTCAAGCACTCGGATACAAACG 785
Db 195 ThrAlaAlaThrThrThrAlaAlaThrThrThrThrAlaAlaThrThrThrThrThrThr 214
QY 786 ACGAAACATCCAGACGCGCTCTCGGCTACCGCGAAACCGTCGAACAGCTCGCCCTCT 845
Db 215 ThrAlaAlaThrThrThrAlaAlaThrThrThrThrAlaAlaThrThrThrThrThrThr 234
QY 846 ACCAAAAATACAGACGGGACGCTACGACTGGCAGAGCGTCCAAACCCGCTCATCAGCG 905
Db 235 ThrAlaAlaThrThrThrAlaAlaThrThrThrSerSerAlaThrThrThrThrThrThr 254
QY 906 ACACCCCTGAAAGGACTCGACGCGGACCGCGCGCAACCCCGGATTTCGGGAGGCTGC 965
Db 255 AlaAlaThrThrThrAlaAlaThrThrThrThrAlaAlaThrThrThr-----AlaAla 271
```

QY 966 AAGACGGCGCTCAAAACACGCCGAGAAAAGCGTCTATCTGTGTTTCAACCGTATTTTCGTCCTCA 1025
Db 272 ThrThrAlaAlaThrThrAlaAlaThrThrAlaAlaThrThrAlaAlaThrThrAlaAla 291
QY 1026 CAAAATCCGCGACAGACGCACTGGCAAAAACCTGGTGAGGACGCATAGAGCTTACCGTCC 1085
Db 292 ThrThrAlaAlaThr-----ThrThrAlaAlaThrThrAlaAlaThrThrAla 305
QY 1086 TGACCAACTCGCTACAGCGCACCGACGTTTCCGCCCGCTCCATTCGGGTACGTCACAAATACC 1145
Db 306 AlaThrThrThrAlaAlaThrThrAlaAlaThrThrAlaAlaThrThrAlaAlaThrThrAla 325
QY 1146 GAAACCGCTGCTCAAAAGCGCGCATCAAACTCTACGAGCTGCAACCCAAACCATCGCGTCC 1205
Db 326 AlaThrThrThrGlySerProThrSerGlySerThrSerThrThrGlyAlaSerThrSer 345
QY 1206 CCGCCAAAGACAAAGGCGCTGACCGGAGCTCCGTAACACGAGCTGCATGCCCAAAACCT 1265
Db 346 ThrProSerAlaSerThrAlaThrSerAlaThrPro----- 357
QY 1266 TCATTGTGGACGGCAAAACGATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGCGAC 1325
Db 358 -----ThrSerThrSerAlaAlaThrThrSerThrProThrProThr 374
QY 1326 GGCTCAATACGAAATGGGCGTGCTCATCGAAAGCCCAAAATCGCAGAACACATGGAGC 1385
Db 375 SerAlaAlaThrSerAlaGluSerThrGluAlaProThrSerThr----- 390
QY 1386 GCACCTCGCGGATACACACCGAATACGCTACCGCTTACCCCTCGACAAACACAAACC 1445
Db 391 -----ProThrThrAspThrThrThrProSerGluAlaThrThr 403
QY 1446 GCCTGCAATGGCACGATCCGCCACCGCAAAACCTACC----- 1484
Db 404 AlaThrThrSer-----ProGluSerThrThrValSerAlaSerThrThrSer 419
QY 1485 -----CGAACGAAACCGAAGCAAACTTTGAAAC 1514
Db 420 AlaThrThrAlaPheThrThrGluSerHisThrSerProAspSerSerThrGlySer 439
QY 1515 -----GCATCGCGCAAAATCTATCCCTGCTGCCCATCG 1550
Db 440 ThrSerThrAlaGluProSerSerThrPheThrLeuThrProSer 454
RESULT 20
T45463
membrane glycoprotein [imported] - equine herpesvirus 1
C:Species: equine herpesvirus 1
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45463
R:Kirisawa, R.; Kobayashi, T.; Kawakami, Y.; Iwai, H.
J. Equine Sci. 7, 79-87, 1996
A:Title: Nucleotide sequences of open reading frames 1, 24 and 71 of an attenuated equin
A:Reference number: Z22973
A:Accession: T45463
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-867 <KIR>
A:Cross-references: UNIPROT:O39782; UNIPARC:UPI000000EF109; EMBL:D88734; PIDN:BA20038.1
A:Experimental source: isolate 3f clone; strain BK343
C:Genetics:
A:Note: ORF71
A:Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homod

Qy	126	GC	GCTCATTTCAATACATCCAAACTGTCCTCTCGAGACAACATCCTTGCAAATCCGGCAC	185
Db	19	A	lailegSerThrThrThrGluThrThr-SerSerSerSerThrSerGlySer	38
Qy	186	CC	CCTCATACAACGGGCTATCCGACATCTACCTGTCTGACGACGCCCCACGAAGCCCTTG	245
Db	39	G	lycInSerThrSerSerGlyThrThrAsnSerSerSerProThrThrSerPro---	57
Qy	246	CC	GGCCCGCGCGCCCTTATCGAATCTTCGGAACACAGCCTCGAATTGCAATATACATTT	305
Db	58	-	-----ProThrThrSerSerSerProProThr-SerThrHisThrSerSerProSer	74
Qy	306	GG	CGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGTACCTTGC CGCAGAAGCG	365
Db	75	S	erThrSerThrGlnSerSerThrAlaAlaThrSerSerSerSerAlaProSerThrAla	94
Qy	366	GC	TGCGCGTAGCGCTCTGTTGGAGACAACAACACGCGCGGGTTGGAGCATCTCCTGC	425
Db	94	-	-----	94
Qy	426	TC	GGCCTCGACAGCCATCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCTACGCA	485
Db	95	S	erSerThrThrSeriLeProThrSer-----ThrSerThrGluThrThrThrThr	111
Qy	486	AA	TGGCGCGCACCTCGGCTACTCTGACG-----ACTTCCC CGCTCAACCGCGCATGC	539
Db	112	T	hrProThrAlaSerThrThrThrProThrThrThrAlaAlaProThrThrAlaAla	131
Qy	540	AC	AAATAATCCTTTACCGCGCAGAACCGCGCCACCATPACTCGCGGAGCAATATCGCG	599
Db	132	T	hrThrAlaVal-----ThrThrAlaAlaSerThrSerAlaGluThrThrThrAla	149
Qy	600	AC	GAATCTTCAAGTTCGGTGAGGACACCGTTTTTCGCGACCTGGACATCCTCGCCACG	659
Db	150	T	hrAlaThrAlaThrSerThrProThrThrThrThrProThrSerThrThrThrThrThr	169
Qy	660	GC	AGCTGTGCGCGAAGTATCGCAGACGACTTCGACCGCTACTCGGCAAGCCATTCGCGCC	719
Db	170	A	laThrThrThrValPro---ThrThrAlaSerThrThrThrAspThrThrThrAlaAla	188
Qy	720	ACA	AGCCGCGCATCATCCGAGCGGGCAACATCGGCAAGGTCITTCAGCACTCGGAT	779
Db	189	T	hrThrThrAlaAlaThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAla	208
Qy	780	ACA	AGCAGAAAACATCCAGACGCGCTCTCGGCTACCGCGAAACCGTCGAACAGTCGC	839
Db	209	T	hrThrThrAlaAlaThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAla	228
Qy	840	CC	TCTACCAAAAAATACAGACGGGACGATCGACTGGCAGAGGTCCTCAACCGCCTGA	899
Db	229	T	hrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAla	248
Qy	900	TC	AGCAGACCCCTGCAAAAGGACTCGACCGCGACCGCGCA-----AACCGCGGATTG	953
Db	249	T	hrThrThrAlaAlaThrThrSerSerAlaThrThrThrAlaAlaThrThrThrAlaAlaThr	268
Qy	954	CC	GGGAGGCTGCAAGACGCGCTCAACAGCGCCGCAAAAAAGCGTCTATCTGGTTTACCCCT	1013
Db	269	T	hrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThr	288
Qy	1014	ATT	TGTCCTACAAAATCCGGCNACAGCAGCATGGCAAAACTGGTCGAGGACGGCATAG	1073
Db	289	T	hrThrAlaAlaThrThrThrAlaAlaThr-----ThrThrAlaAla	302
Qy	1074	AC	GTTACGTCCTGACCATTCGCTACAGCGCAGCCAGTTCGCGCGTGCATTCCCGGCT	1133
Db	303	T	hrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAla	322
Qy	1134	AC	GTCAAATACCGAAAAACCGCTGCTCAAGCGCGCATCAAACTCTACGACTGCAACCCA	1193
Db	323	T	hrThrThrAlaAlaThrThrThrGlySerProThrSerGlySerThrSerThrThrGly	342

US-10-665-990A-13 (1-1561) x T45463 (1-867)

QY	221	CTCGACGACCCCGCCAGAAAGCCCTTGC	CGCCGCGCGCCCTTATCGAATCTCCCGAACAC	280
		:::	:::	
Db	138	LeuThrAsnGlyGluGluThrPheProValLeuLeuAspAlaLeuArgylsAlaGluAsn	157	
		:::	:::	
QY	281	AGCCTCGAATTGCCAATACATATTGGCGCAACGACATTTCCGCGAGGCTGCTGTTCAAC	340	
		:::	:::	
Db	158	HisIleHisIleGlnTyrIlePheLysThrAspAlaIleSerThrGluIleArgAsp	177	
		:::	:::	
QY	341	CTCATGTACCTTGGCGCAGAACGGCGCTGCGCTACGCCTGCTGTGGACGACACCAAC	400	
		:::	:::	
Db	178	IleLeuValGluLysAlaLysSerGlyValGluValArgPheMetPheAspGlyLeuGly	197	
		:::	:::	
QY	401	ACGCGCGGTTGGACGATCTCTCTGCTCGCCCTCGACGCCATCCCAATATCAAGTGC	460	
		:::	:::	
Db	198	SerSerLysLeuGlyLysAlaPheLeuAlaProLeuLysGluAlaGlyValSerIleHis	217	
		:::	:::	
QY	461	CTGTTCAACCCCTTCGTCTACGCAAAATGGCGCGCACTCGGCTACCTGACCGACTTCC	520	
		:::	:::	
Db	218	AlaPheAspPro--IleAlaSerProIrrIleVal-----Arg	229	
		:::	:::	
QY	521	CGCCTCAACCGCGATGCACAAACAATCTTTACCGCGGACAAACGGCGCACCATCTC	580	
		:::	:::	
Db	230	ThrAlaAsnLeuArgAsnHisArgLysIleValIleAspGlyGlnIleGlyPheThr	249	
		:::	:::	
QY	581	GGCGGACGCAATATCGGCGCAGCAATAC-----TTCAAACTCGGTGAG	622	
		:::	:::	
Db	250	GlyGlyLeuAsnIleGlyGluGluTyrArgSerAsnThrProAspPheArgValTrpArg	269	
		:::	:::	

977	AAACAGCCGCAAAAAGCGTCTATCTGGTTTCAACCTATTTTCGTCCTCACAACATCCGCG	103
342	-----MetLeuAspLeuIle	344
Qy	
Db	
977	AAACAGCCGCAAAAAGCGTCTATCTGGTTTCAACCTATTTTCGTCCTCACAACATCCGCG	103
342	-----MetLeuAspLeuIle	344
Qy	
Db	
347	AspSerAlaLysGluSerValTrpIleValSerProTyrPheValProAspGluGluSer	366
Qy	
Db	
1037	ACAGACGCCTGGCAAAACTGGTGCAGACGCGATAGACGTTACCGTCTCTGACCAACTCG	1099
367	LeuAlaValIleArgAvalAlaMetSerGlyValAspAlaGValIleIleProGly	386
Qy	
Db	
1097	CTACAGCGCAGCGAGTTCGCGCGTCCATTTCGCGCTACGTCAAATACCGAAAAACCGCTG	1155
387	-----LysGlyAspArgGlyIleSerPheHisGlySerAsnAlaTyrValIysThrMet	404
Qy	
Db	
1157	CTCAAGCCGGCATCAAACTCTACGAGCTGCAACCCCAACCATCGCGTCCCGCCACAAAA	1215
405	IleGluAlaGlyAlaLysMetTyrAlaAlaAspSerPheVal	420
Qy	
Db	
1217	GACAAAGGCCTGACCGCGAGCTCGTAAACGACCTGCATGCCAAAACTTCATTGTGTGAC	1276
421	-----HisAlaLysAlaMetLeuValAsp	428
Qy	
Db	

```
QY 1277 GCGAAGCGACTTTCATCGGCTCATTAACCTCGACCCCGCTTCGCGACGGCTCAATACC 1336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyThrArgAlaAlaileGlyThrAlaAsnPheAspValArgSerPheArgLeuAsnHis 448

QY 1337 GAAATGGCGTGTGTCATCGAAGCCCAAAATCGGAGAACACAGATGGAGCGCACCCCTCGCC 1396
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GluLeuMetValPheLeuTyArgPhe-----SerGluAlaMetHisLeuLysArg 466

QY 1397 GATACACACCGCAATAGCCCTACCGCTTACCCTCGACAAACACACACCGCTGCAATGG 1456
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AspPheLysLysAspPhe-----GluAspSerArgLeu----- 477

QY 1457 CACGATCCGCGCCCGAAGAAACCTACCGACGACCGAAGCCAAACTTGGAAACGC 1516
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db PheThrMetLysAspMetGluAsnLysProLeuLeuThrArgIleLysGlu 494

QY 1517 ATCGCGCGAAMAACTCTATCC 1537
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ValLeuSerSerLeuLeuSer 501

RESULT 22
A:lin2646
cardiolipin synthase homolog lin2646 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11762
R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluecker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
C:Accession: A11762
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <GLA>
A:Cross-references: UNIPROT:Q92720; UNIPARC:UPI00000CC933; GB:AL592022; PIDN:CAC97873.1;
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2646
C:Superfamily: cardiolipin synthase
```

```
Alignment Scores:
Pred. No.: 1.13e-09 Length: 482
Score: 292.00 Matches: 109
Percent Similarity: 39.0% Conservative: 56
Best Local Similarity: 25.8% Mismatches: 164
Query Match: 10.2% Indels: 94
DB: Gaps: 15
```

US-10-665-990A-13 (1-1561) x A11762 (1-482)

```
QY 206 TCCGACATCTACTGCTCGACACCCCGACGAGCCCTTGGCGCGCGCCCTTATC 265
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AsnGluValGluLeuPheIleAspGlyHisGluLysPheAspAlaLeuIleAlaAspIle 138

QY 266 GAATCTGCGGACACACAGCTCGATTTGCAATACTATTGCGCAACGACATTTCCGGC 325
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GluLysAlaLysAspHisIleHisLeuIleTyTyIlePheHisSerAspGluLeuGly 158

QY 326 AGGCTGTGTTCAACTCATCTACCTTGGCGGAGAACCGCGGTGGGTAGCCTGTG 385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AsnArgLeuMetArgValLeuGluArgLysAlaAlaGluGlyLeuAsnValLysIleIle 178

QY 386 TTGGACGACAAACACGCGCGGGTTGGACGATCTCTCTCGCCCTCGACGCCATCCC 445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db TyrAspAlaMetGlySerArgThrThrLysLysSerPhePheArgThrPheGluLysAsn 198

QY 446 AATATCGAAGTGGCGGTGTTCAACCCCTTCGTCCTACGCAAAATGGCGGCACTCGGCTAC 505
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyGlyLeuValArgProPhePheProSerLysLeuPro----- 211
```

1018 1019 1078 1138 1139 1198 1199 1258 1318 1319 1378 1414 1415 461 462 464

CTGACCGACTTCCCGCGCTCAACCGCGCGCATGCACAAATAATCCTTTACCGCGCACAC 565
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LeuIleAsnPhe---ArgLeuAsnTyArgAsnHisArgLysLeuAlaIleAAspGly 230

566 CGCGCCACCATCTCGCGGACGCAATATCGGCGACGAATCTTTCAAAGTCGGTGAGGAC 625
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AspValGlyTyriIleGlyGlyPheAsnIleGlyAspGluTy---LeuGlyArgSer 248

626 ACCGTTTTTCGCGCACCTGGACATCTCGCCACCGCGACGCTGCTGGCGAAGTATCGCAC 685
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LysLysPheGly----- 252

686 GACTTCGACCGCTACTGGGCAAGCCATTCCGCCACACGCGCGCATCATCGCGACG 745
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ---TyrTrpArgAsp-----ThrHisLeuArgValHis 262

746 GGCAACATCGGCAAGGTCTTCAAGCA-----CTCGGATACAAACGACGCAACATCC 796
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyLysAlaValTyAlaMetGlnThrArgPheIleMetAspTrpAsnSerAlaSerSer 282

797 AGACACGCGCTC-----CTCGCTACCGCGAAACCGTCTC-----GAACAGTCGCGCC 841
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ThrHisLysIleAspTyLysAlaArgTyPheProThrPheHisGlyLysGlyHisThr 302

842 CTCTACCAAAATAATACAGCGGA---CGCATCGACTGGCGAGAGCGTCCAAACCGCGCTG 898
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db SerMetGlnIleValSerSerGlyProAspSerGlnTrpGlnGlnIleLysAsnGlyTy 322

899 ATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAAAACCGCGGATTCGCCGG 958
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db IleLys----- 324

959 AGGCTGCAAGACGCGCTCAAAACAGCCCGGAAAGCGTCTATCTGTTTCACTTATTC 1018
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db IleProAspAlaSerLeuLeuGluAlaIleLysIleAlaAlaLeuSerGlyValAspVal 360

1079 ACCGTCCTGACCAACTCGCTACAGCGACCGCGTTCGCGCGCTCCATTCGCGGTACGTC 1138
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ArgValMetIleProAsn-----LysProAspHisAlaPheValTyArgAlaThrThr 378

1139 AAATACCGAAACCGCTGCTCAAAAGCGCGCATCAAACTCTACGAGCTCCAAACCAACCAT 1198
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AsnTyAlaGlyGluLeuMetGluThrGlyAlaLysIlePhe----- 392

1199 GCGTCCCGCGCACAAAGACAAAGCGCTGACCGCGAGCTCCGTAACAGCGCTGCATGCC 1258
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ---IleTyAspAsnGlyPhe-----IleHisAla 401

1259 AAAACCTTCATTGTGGCGCAAAACCGCATCTTCATCGGCTCATCACTCGACCCCGCT 1318
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LysThrLeuValValAspGlyGluIleAlaSerValGlyThrAlaAsnMetAspPheArg 421

1319 TCCGACCGCTCAATACCGAAATGGCGTCTCGAAAGCCCGCAAAATCGCAGAACAG 1378
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db SerPheArgLeuAsnPheGluValAsnAlaPheIleTyGluLysLysMetValGlnLys 441

1379 ATGGAGCGCACCTCGCCGAT-----ACCACACCCGAA---TAC 1414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LeuGluAspAlaPheLeuGluAspIleLeuLysSerTyGlnLeuThrProGluLeuTy 461

1415 GCCTACCGC 1423
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AlalysArg 464

RESULT 23

AG1387

cardiolipin synthase homolog lmo2503 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1387

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AG1387

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-482 <GLA>

A;Cross-references: UNIPROT:Q8Y4E3; UNIPARC:UPI0000055715; GB:NC_003210; PIDN:CAD00581.1

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo2503

C;Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 2,21e-09 Length: 482
Score: 287.00 Matches: 108
Percent Similarity: 39.0% Conservatives: 57
Best Local Similarity: 25.5% Mismatches: 164
Query Match: 10.1% Indels: 94
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x AG1387 (1-482)

```
Qy 206 TCGACACTCTACCTCGACGACCCCAAGCCCTTGGCCGCGCGCCCTTATC 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 AsnGluValGluLeuPheValAspGlyHisGluLysPheAspAlaLeuAlaAspIle 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 266 GAATCTCGGACACAGCCTCGATTGCAATACATTTGGCGCAAGCACATTTCCGGC 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 GluLysAlaLysAspHisHisLeuIleTyrTyrIlePheHisSerAspGluLeuGly 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 326 AGGCTGTGTTCAACCTCATCTTACCTGCGCAGAACGCGCGTACGCTGCTG 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 AsnArgLeuMetArgValLeuGluArgLysAlaAlaGluGlyLeuAsnValLysIle 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 386 TTGACGACAAACACAGCGCGGTGGACGATCTCTGCTCGCCCTCGACAGCATCCC 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 TyrAspAlaMetGlySerArgThrThrLysLysSerPhePheArgThrPheGlnLysAsn 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 446 AATATCGAAGTGCCTGTTCAACCCCTTCTGCTCTACGCAATGCGCGGCACTCGGTAC 505
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 GlyGlyLeuValArgProPheProSerLysLeuPro----- 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 506 CTGACCGACTTCCCGCGCTCAACCGCGCATGCAACAACAACTTTACCGCGCAAC 565
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 LeuIleAsnPhe---ArgLeuAsnTyrArgAsnHisArgLysLeuAlaIleAspGly 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 566 CGCGCCACCATATCTCGCGGACGCAATATCGGACGCAATCTTCAAAGTGGTGAGGAC 625
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 AspValGlyTyrIleGlyGlyPheAsnIleGlyAspGluTyr-----LeuGlyAlaSer 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 626 ACCGTTTTCGCGACCTGGACATCTCTCGCACCGCGCGCTGTCGGGAGATTCGCAC 685
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 LysLysPheGly----- 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 686 GACTTCGACCGCTACTCGGCAAGCACTTCGCGCCCAACAGCGCGCATCATCCGACG 745
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 -----TyrTrpArgAsp-----ThrHisLeuArgValHis 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 746 GGCACATCGCAAGGTTCTTCAAGCA-----CTCGGATACACGACGAAACATCC 796
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 GlyLysAlaValTyrAlaMetGlnThrArgPheIleMetAspTrpAsnSerAlaSerSer 282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 797 AGACACGGCTC-----CTCGGCTACCGCAACCGTC-----GACAGTCCGCC 841
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 ThrHisLysIleAspTyrLysAlaArgTyrPheProThrPheHisGlyLysGlyHisThr 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Qy 842 CTCCTACCAAAAATAACAGACGGGA---CGCATCGACTGGCAGAGCGTCCAAACCCGCTG 898
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 SerMetGlnIleValSerSerGlyProAspSerGlyTrpGlnGlnIleLysAsnGlyTyr 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 899 ATCAGCGACACCCCTGCAAAAGGACTCGACCGCAGCCGCCGCAAAACCGCGATTGCCGGG 958
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 IleLys----- 324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 959 AGGCTGCAAGACGCGCTCAACACGCCCGAAAAAAGCGTCTATCTGTTTTCACCCATTTC 1018
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 -----MetIleAsnAlaIleLysLysThrIleTyrLeuGlnSerProTyrPhe 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1019 GTCCTCAAAAATCCGCGACACGCGCTCAAAAGCTGGTGCAGGACGGCATAGACGTT 1078
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 341 IleProAspAlaSerLeuLeuGluAlaIleLysIleAlaLeuSerGlyValAspVal 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1079 ACCGTCCTGACCAACTCGCTACAGCGACCGACGCTTCGCGCGCTCCATTCCGGCTACGTC 1138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 ArgValMetIleProAsn-----LysProAspHisAlaPheValTyrArgAlaThrThr 378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1139 AAATACCGAAAACGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCCCAACCAT 1198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 379 AsnTyrAlaGlyGluLeuMetGluThrGlyAlaLysIlePhe----- 392
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1199 GCGTCCCGCCCAACAAAAGGCGCTGACCGGCGAGCTCCGTACCGCCTCGCATGCC 1258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 393 -----IleTyrAspAsnGlyPhe-----IleHisAla 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1259 AAAACCTTCATTGTGGACGCGCAACGCATCTTCATCGGCTCATTCACCTCGACCCCGCT 1318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 LysThrLeuValValAspGlyGluIleAlaSerValGlyThrAlaAsnMetAspPheArg 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1319 TCCGACGGCTCAATACCGAATGGCGCTGTCATCGAAAGCCCAAAATTCGCGAGACAG 1378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 SerPheArgLeuAsnPheGluValAsnAlaPheIleTyrGluLysGlnMetValGlnLys 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1379 ATGAGCGCGCCCTCGCGCAT-----ACCACACCGGAA---TAC 1414
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 LeuGluAspAlaPheLeuGluAspIleLeuLysSerTyrGlnLeuThrProGluLeuTyr 461
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1415 GCCTACCGC 1423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 462 AlaLysArg 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 24
T45025
mucin MUC5B, tracheobronchial [imported] - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T45025
R;Dessens, J.L.; Guyonnet-Duperat, V.; Porchet, N.; Aubert, J.P.; Laine, A.
J. Biol. Chem. 272, 3168-3178, 1997
A;Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alterna
A;Reference number: Z22899; MUID:97166151; PMID:9013550
A;Accession: T45025
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3570 <DES>
A;Cross-references: UNIPARC:UPI000016AD88; EMBL:Z72496; NID:g1834502; PIDN:CAA96577.1;
A;Experimental source: placenta
C;Genetics:
A;Gene: MUC5B
Alignment Scores:
Pred. No.: 2,94e-09 Length: 3570
Score: 286.00 Matches: 155
Percent Similarity: 31.0% Conservatives: 67
Best Local Similarity: 21.7% Mismatches: 211
Query Match: 10.0% Indels: 282
DB: 2 Gaps: 27
US-10-665-990A-13 (1-1561) x T45025 (1-3570)
```

Qy	141	CTTCCAAACCTGTCTCTCTGGACAACAATCCTGCAAAATCCGGCACACCCCTCATAAACA---	197
Db	1868	LeuProAlaLeuArgSerThrAlaThr-----ThrProThrAlaThrSer	1882
Qy	198	--ACGGGCTATCGACATCTACCTGC-----TCGACG	227
Db	1883	ValThrAlaIleProSerSerLeuGlyThrAlaTrpThrArgLeuSerGlnThrThr	1902
Qy	228	ACCCCCAGAACCCCTTGCCGCCGCGCCCTTATCGAATCTGCGC-----AAC	278
Db	1903	ThrProThrAlaThrMetSerThrAlaThrPro---SerSerThrProGluThrValHis	1921
Qy	279	ACAGCTCGATTGCAATACTACATTTTGGCGCAAACGACATTTCCGGCAGCGCTCTGTCCA	338
Db	1922	ThrSerThrValLeuThrThrThr-----AlaThrThrArgThrGlySerValAla	1939
Qy	339	ACCTCATGTACCTTGGCGCAGACGGCGGTGGCTACGCTGCTGTGGACGACAACA	398
Db	1940	ThrProSerSerThrPro-----GlyThrAlaHisThrThrLysValProThrThr	1956
Qy	399	ACAGCGCGGGTGGACGATC-----TCCTGTCTGCCTTCGACAGCATC-----	443
Db	1957	ThrThrThrGlyPheThrAlaThrProSerSerSerProGlyThrAlaLeuThrProPro	1976
Qy	444	-----CCAAATATCGAAGTGGCCCTGTTTCAACCCCTTCGTCTCC	479
Db	1977	ValTpIleSerThrThrThrProThrThrArgGlySerThrValThrProSerSer	1996
Qy	480	TACGCAATGGCGGCACTCGGTACCTGACCGACTCCC CGCCTCAACCGCGCA---	536
Db	1997	IleProGlyThrThrHisThrAlaThrValLeuThrThrThrThrThrValAlaThr	2016
Qy	537	-----TGCAACAACAAATCTTTACGCGC-----	560
Db	2017	GlySerMetAlaThrProSerSerSerThrGlnThrSerGlyThrProProSerLeuThr	2036
Qy	561	ACAACCGCGCACATACTCGGGGACGCAATATCGCGACGAATACTTCAAAGTCGGTG	620
Db	2037	ThrThrAlaThrThrIleThrAlaThrGlySerThrThrAsnProSerSerThrProGly	2056
Qy	621	AGGACACGTTTTTCGCG- ----ACCTGGACATCTCTCGCACCGGACGCTCGTCGGCG	674
Db	2057	ThrThrProIleProProValLeuThrThrAlaThrThrProAlaAlaThrSerSer	2076
Qy	675	AAGTAT-----CGCACGACT	689
Db	2077	ThrValThrProSerSerAlaLeuGlyThrThrHisThrProProValProAsnThrThr	2096
Qy	690	TCGACCGCTACTGGGAAGCATTCCGCCCAACAGCCACGCGCATCATCCGAGCGGCA	749
Db	2097	AlaThrThrHisGlyArgSerLeuProProSerSerProHisThrValProThrAlaTrp	2116
Qy	750	ACATCGCAAGGGTC-----TTCAAGCACTCGGATACAACGACG	788
Db	2117	ThrSerAlaThrSerGlyIleLeuGlyThrThrHisIleThrGluProSerThrGlyThr	2136
Qy	789	AAACATCCAGACACGCGCTCTGCGCTACCGCGAAACCG-----TCGAAC	833
Db	2137	SerHisThrProAlaAlaThrThrGlyThrThrGlnProSerThrProAlaLeuSerSer	2156
Qy	834	AGTCGCCCTCTACCAAAAATAACAGAGGGACGCATCGACTGGCAGAGCGTCCAAACC	893
Db	2157	ProHisProSerSerArgThrThrGluSerProProSerProGly-----	2171
Qy	894	GCCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGCGCGCAACCGCGCATTG	953
Db	2172	-----ThrThrThrProGlyHisThrArgGlyThrSerArgThrThrAlaThrAlaThr	2189
Qy	954	CCGGAGGCTGCAAGACGCGCTCAACAGCCCGCAAAAAAGCGTCTATCTCGTTTCACCT	1013
Db	2190	ProSer-----LysThrArgThrSerThrLeuLeuProSerProThrSerAlaPro	2207

QY	1014	ATTTCGCTCCCTACAAATCCGGCA-----CAGACGCACGTGGCAAACACTCG-----	1058
Db	2208	IleThrThrValThrThrGlyCysGluProGlnCysAlaTrpSerGluTrpLeuAsp	2227
QY	1058	-----	1058
DB	2228	TyrSerTyrProMetProGlyProSerGlyGlyAspPheAspThrTyrSerAenIleArg	2247
QY	1059	-----TGcAGGAcGGCATAGAcG	1076
Db	2248	AlaAlaGlyGlyAlaValCysGluGlnProLeuGlyLeuGluCysArg--AlaGlnAla	2266
QY	1077	TTACCGTCCTGACCACCTCGCTACAGCGCACCGACGTGGCCG-----	1118
Db	2267	GlnProGlyValProLeuArgGluLeuGlyGlnValValGluCysSerLeuAspPheGly	2286
QY	1118	-----	1118
Db	2287	LeuValCysArgAsnArgGluGlnValValGlyLysPheLysMetCysPheAsnTyrGluIle	2306
QY	1119	-----CGTCcATTCCGGCTAcGTCAAATAcCGAA	1148
Db	2307	ArgValPheCysCysAsnTyrGlyHisCysProSerThrProAlaThrSerSerThrAla	2326
QY	1149	AACCGCTGCTCAAGCGCGCATCA-----	1172
Db	2327	ThrProSerSerThrProGlyThrThrTyrTrpIleLeuThrLysLeuThrThrAlaThr	2346
QY	1172	-----	1172
Db	2347	ThrThrGluSerThrGlySerThrAlaThrProSerSerThrGlnGlyProProAlaGly	2366
QY	1173	-----AACTcTAGAcGTGAACCCAAcCATGCCG-----	1202
Db	2367	ThrProHisValSerThrThrAlaThrThrProThrValThrSerSerLysAlaThrPro	2386
QY	1203	---TCCCcCGCAcAAAAGAcGGCTGAcCGGcAGCTCCGTAACAcGAGcCTGCATGCCCA	1259
Db	2387	PheSerSerProGlyThrAlaThrAlaLeuProAlaLeuArgSerThrAlaThrThrPro	2406
QY	1260	AAAcCTTcATTGGAcGGCAAAcCGAcATCTTCATCG-----	1295
Db	2407	ThrAlaThrSerPheThrAlaIleProSerSerSerLeuGlyThrThrTrpThrArgLeu	2426
QY	1295	-----	1295
Db	2427	SerGlnThrThrThrProMetAlaThrMetSerThrAlaThrProSerSerThrProGlu	2446
QY	1295	-----	1295
Db	2447	ThrValHisThrSerThrValLeuThrThrThrAlaThrThrThrGlyAlaThrGlySer	2466
QY	1296	-----GCTcATTCAcCTCGAcCCCC-----	1316
Db	2467	ValAlaThrProSerSerThrProGlyThrAlaHisThrThrLysValProThrThrThr	2486
QY	1317	-----GTTCCGCAcCGGTCA-----ATAcCGAAAA	1340
Db	2487	ThrThrGlyPheThrValThrProSerSerSerProGlyThrAlaargThrProProVal	2506
QY	1341	TGGCGcGTGCATCGAAAGCCCAAAATCCGAGAACAAGATGGAGCGCACCTCGCCGATA	1400
Db	2507	TriPileSerThrThrThrProThrProThrSerGlySerThrValThrProSerSerIle	2526
QY	1401	CCA---CACCCGAATAcCGCTAcCGGTtACCTTCGACAAACAACAcCGCTGCATGGC	1457
Db	2527	ProGlyThrThrHisThrProThrValLeuThrThrThrThrGlnProValAlaThrGly	2546
QY	1458	AcGATCCCGCCAcCCGAAAAAcCTAcCCGAAcCGAACCCGAAGCCA	1502
Db	2547	SerMetAlaThrProSerSerSerThrGlnThrSerGlyThrThrPro	2561

[illegible]

RESULT 26

Alignment Scores:

Pred. No.: 4,06e-09 Length: 481
Score: 282.50 Matches: 104
Percent Similarity: 39.0% Conservativity: 76
Best Local Similarity: 22.5% Mismatches: 167
Query Match: 9.9% Indels: 115
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x H86744 (1-481)

```
QY 194 ACAACGGGCTATCCGACATCTACTGCTCGACGACCCCAACGAGCCCTTCCCGCCGC 253
DB 118 AsnThrGly-----VallyIlePheAsnAspGlyArgGlnLysPheAsnAlaLeu 134
QY 254 GCGGCCCTTATCGAATCTCCGAAACACAGCCTCGATTGCAATCTACATTTGGCGCAAC 313
DB 135 IleAspAspIleAsnLysAlaLysHisHisValHisMetGluTyrTyrIlePheArgMet 154
QY 314 GACATTTCCGGAGGCTGCTGTTCACCTCATGTATCTTTCGCGAGAACGGCGCTGCGC 373
DB 155 AspArgMetGlyHisGluIleTyrAspAlaLeuLeuAlaAlaLysArgGlyValGlu 174
QY 374 GTACGCTCTCTGTGGAC-----GACACACACACGCGCGGTGGACGATCTCTG 424
DB 175 ValLysLeuLeuIleAspAlaTyrGlySerAsnLysThrLys----- 188
QY 425 CTCGCCCTCGACAGCCATCCCAATATATCGAAGTGGCGCTGTTCACACCCCTTCGCTACGC 484
DB 189 -----MetSerAsnPhelysGluLeuGln----- 196
QY 485 AAATGGCGGCACTCGGTACTCGACCGACTTC----- 517
DB 197 -----GluAlaGlyHisValAlaGlnPhePheProLeuIleMetProLeuValAsn 214
QY 518 CCGCGCTCAACGGCGCATGCACACAAATCTTTACCGCGCAACCGCGCCACCAT 577
DB 215 ProArgThrAsnTyrArgLeuHisArgLysIleValIleAspGlyIleleGlyTyr 234
QY 578 CTCGGCGGACGAATATCCGGACGAATACTTCAAAGTCGTGAGGAC---ACCGTTTTC 634
DB 235 ThrGlyGlyPheAsnValGlyAspGluTyrAlaSerIleThrLysLysPheGlyTyrTrp 254
QY 635 GCGGACTTGGACATCTCCGACCGGACGCTGTCGGGGAAGTATCGCACTTCGAC 694
DB 255 ArgAspAsnHisLeuArgLeuThrGlyAspIleValTyrSerLeuGlnHisArgPheIle 274
QY 695 CGCTACTGGCAAGCATTCGCCCCACACGCGCATCATCCGACGCGCAACATC 754
DB 275 LeuAspTrp----- 277
QY 755 GGCAAGGGTCTTCAAGCACTCGGATACACGACGAAACATCCAGACGCGCTCAACAGCCGCG 814
DB 278 -----AsnSerGlnHis----- 281
QY 815 TACCGGAAACCGTCGACAGTCGCCCTCTAC---CAAAAATACAGACGGGACGATC 871
DB 282 HisPheGluValThrGluGlyGluProTyrPheProAspSerIleValGluGlyHis--- 300
QY 872 GACTGCGACAGGTCACAAACCGCTGATCAGCGACACCCCTGCAAAAGACTCGACGCG 931
DB 301 -----ValAlaThrGlnLeuValThrSerGlyPro----- 310
QY 932 GACCGCGCAAAACCGCGGATTCGGGGAGGCTGCAAGACGCGCTCAACAGCCGCAAAA 991
DB 311 AspGluAspLysGluGlnIleLysLeuThrTyrMetLysMetIleSerGlyAlaGluArg 330
QY 992 AGGCTATCTGTTTCAACCTATTTCGTCCTTACAAATCCGGACACAGCAGCAGCTGGA 1051
DB 331 GluIleIleIleGlnThrProTyrTyrIleProSerAspAlaLeuHisGluSerIleLys 350
QY 1052 AAACGTGTCGACGCGCATAGACGTTACCGTCTCGACCACTCGCTACAGCGCGCGAC 1111
```

```
DB 351 LeuAlaLeuLeuSerGlyValGlnValLysLeuLeuIleProAsn-----LysProAsp 368
QY 1112 GTTCCGCGCGTCCATTTCCGGCTACGTCAAATACGAAACCGCTCTCTCAAGCCGGCATC 1171
DB 369 HisProLeuValTyrTrpAlaThrTyrPheHisAlaAlaAspLeuValLysTyrGlyAla 388
QY 1172 AAACCTCTAGAGCTGCAACCCACCATCGCTCCCGCCACAAAAGACAAAGGCTGACC 1231
DB 389 LysValTyr-----ThrTyrGluAsnGlyPhe--- 397
QY 1232 GGCAGCTCCGTAACCGAGCTGTCACCAACCTTTCATTGTGGAGCGCAACGCACTTTC 1291
DB 398 -----ValHisSerLysThrLeuIleIleAspGlyGluTyrAlaSer 411
QY 1292 ATCGGCTCATTTCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGGCGTCTG 1351
DB 412 ValGlySerAlaAsnLeuAspTyrArgSerLeuGlnLeuCysPheGluAlaAsnValVal 431
QY 1352 ATCGAAACCCCAAAATCGCAGACAGATGAGGCGCACCTTCGCGGATACCACACCCGAA 1411
DB 432 IleTyrAspTyrAspIleSerGlnLysLeuArgAsn----- 443
QY 1412 TACGCTTACCGGTTTACCTTCGACAAACACACACCGCTGCAATGSCACGATCCCGCCACC 1471
DB 444 -----AspPheMetLysAspLeuLysLeuSerArgProLeuThr 456
QY 1472 CGAAAAACCTTACCGAAGCCGCAAGCCAAACTTTGGAAACCGCATCCGCGCAAAATC 1531
DB 457 LeuGluArgTyrGluGluArgSerLysLeuValArgPheLysGluGlyLeuAlaArgLeu 476
QY 1532 CTATCC 1537
DB 477 IleAla 478
```

RESULT 27

B84007

cardiolipin synthetase B82858 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: B84007

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: B84007

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-503 <STO>

A:Cross-references: UNIPROT:Q9K824; UNIPARC:UPI0000127B9A; GB:AP001516; GB:BA000004; NII

A:Experimental source: strain C-125

C:Genetics:

A:Gene: B82858

C:Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.: 9.81e-09 Length: 503
Score: 276.00 Matches: 110
Percent Similarity: 37.2% Conservativity: 57
Best Local Similarity: 24.5% Mismatches: 188
Query Match: 9.7% Indels: 94
DB: 2 Gaps: 12

US-10-665-990A-13 (1-1561) x B84007 (1-503)

```
QY 167 ATCTGCAATCCCGGCACACCCCTCATTAACACCGGCTATCC-----GACATCTAC 217
DB 124 LeuPheArgLeuAlaHisArgLeuAlaAsnAsnProIleSerPheSerThrAsnThrLys 143
QY 218 CTGCTCGACGACCCCGACGAGCCCTTCGCCCGCGCGCCCTTATGAAATCTCGCGAA 277
DB 144 ValLeuThrAspGlyLysGluThrPheAlaHisIleLysGlnAlaLeuArgMetAlaThr 163
QY 278 CACAGCCTCGATTTCGCAATACTACATTGCGGCAACGACATTTCCGCGGCGGCTGTTTC 337
```



```
Db 178 AspTyrHisLeuArgLeuThrGlyAspGlyValGlnAspLeuGlnAspGlnPheIleGln 197
QY TACTGGCAAGCATTCGCCGCACAAACCCACGCGCATCATCCGCGAGCGCAACATCGGC 757
Db 198 AspTrpGluArgGluSerArgLeuProValThrArg----- 209
QY 758 AAGGTCTTCAAGCACTCGGATACAAACGACGAACATCCAGACGCGCTCTCGGCTAC 817
Db 209 ----- 209
QY 818 CGCGAACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGCGGACGATCGACTGG 877
Db 209 ----- 209
QY 878 CAGAGCGTCCAAACCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGGACGCG 937
Db 210 ----- 210
QY 938 CGCAACCGCGATTGCCGGAGGCTGCAAGCGG-----CTCAACAGAGCC 985
Db 226 IleIleProThrAsnGlySerPheLeuGluAspSerPheLeuGlnLeuValGluGlnAla 245
QY 986 GAAAAAGCGTCTATCTGTTTCACCCCTATTTCGCTTACAAATCCGACGCGCACAGAGCA 1045
Db 246 GluGluThrIleThrIleGlyThrProTyrPheIleProGlyGlyLeuHisAla 265
QY 1046 CTGCAAACTGGTCAGGCGGATACGAGCTTACCGTCTGACCAACTCGCTACAGCG 1105
Db 266 LeuLeuAspAlaAlaArgGlyValAlaValArgLeuLeuValPro-----LysLys 283
QY 1106 ACCGAGTTGGCGGCTGATTCGCGGTACGTCAAAATACCGAAACCGCTGCTCAAGCC 1165
Db 284 GlyAspHisProLeuValLysGluAlaAlaPheProTyrPheLysGluLeuGluGly 303
QY 1166 GGCATCAAACTCTAGAGGTGCAACCCACCATGCGGTCGCCGCCACAAAAGAGAAAGGC 1225
Db 304 GlyIleAsnIleTyrArgTyr----- 313
QY 1226 CTGACCGGAGCTCGGTAAACAGCGCTGATGCAAAACCTTCAATTGTGGAGCGCAACGC 1285
Db 314 Phe-----PheHisGlyLysAlaIleValIleAspLysLeu 326
QY 1286 ATCTTCATCGGCTCATTTCACTCAGCCCGCTTCGCGACGCTCAATACCGAATGGCG 1345
Db 327 AlaAspValGlyThrAlaAsnPheAspLysArgSerPheArgLeuAsnTyrGluIleAsn 346
QY 1346 GTCGTCATCGAAAGCCCAAAATCGCAAGACAGATGAGCGCACCTCGCGGATACCACA 1405
Db 347 CysLeuLeuTyrAspLysGluMetIleGlnLeuValArgGluGluLeu----- 362
QY 1406 CCCGAATACGCTACCGGTTTACCCTCGACAAACAAACCGCTGCAATGGCAGCATCCC 1465
Db 363 ---AspTyrAspPheSerIleSer-----GluArgLeuGlnMetGluAspLeu 377
QY 1466 GCCACCCGAAAAACCTACCGGACCAACCCGAA 1498
Db 378 AlaAsnArgSerPhePheHisArgThrLysGlu 388
```

RESULT 29

T43863
cardiolipin synthase [imported] - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43863
R:Koyama, M.
submitted to the EMBL Data Library, August 1998
A:Description: Clostridium perfringens hem operon.
A:Reference number: Z22707
A:Accession: T43863
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-476 <KOY>

A:Cross-references: UNIPROT:Q9ZNC6; UNIPARC:UPI000016EODF; EMBL:AB017186; PIDN:BAA74786.
A:Experimental source: strain NCTC8237; vegetative cell
C:Genetics:
A:Gene: cls
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 1.46e-08 Length: 476
Score: 273.00 Matches: 108
Percent Similarity: 42.1% Conservative: 79
Best Local Similarity: 24.3% Mismatches: 167
Query Match: 9.6% Indels: 90
DB: 2 Gaps: 20

US-10-665-990A-13 (1-1561) x T43863 (1-476)

```
QY 107 CTGGAAGAACCGACGCGAAAGCCGTCATTTCATATCAATCTTCCAAACCT---GTCCTCTCTGCAC 163
Db 66 LeuAspAspLysThrLysGlnLysTyrLeuAsnSerPheLysSerHisTyrLysLeuAsp 85
QY 164 AACATC-----CTGCAATCCGGCACACCCCTCATCAACACGCGGCTA 205
Db 86 AsnIleSerLeuLysTyrLysAspLeuIleMetMetAsnPheAsnAsnAspAsnSerThr 105
QY 206 -----TCGCATCTACTGCTCGACGACCCCGACGAGCCCTTGCCTCCCGC 253
Db 106 TyrThrGlnArgAsnAspIleAspLeuTyrPheAspAlaAsnSerLeuPheGluMet 125
QY 254 GCCGCCCTTATCGAATCTGCGGAACACACGCTCGAATTGCAATATCATATTGGCGCAAC 313
Db 126 IleAspGluIleAsnLysAlaGluLysPheIleHisMetGluPheTyrIlePheLysSer 145
QY 314 GACATTTCCGCGACGCTGCTGTTCAACTCATGTATCTTGCCTGCGGAGAACGCGGCTGCGC 373
Db 146 AspGluIleGlyLysLysIleLeuGlnAlaLeuThrLysLysAlaLysGluGlyValGlu 165
QY 374 GTACGCTGCTGTGGAGGAC-----AACAAACGCGCGGTTGGAC---GATCTCTCTG 424
Db 166 ValLysLeuLeuValAspSerIleGlyAsnSerIleHisLysLysAspIleAspLysLeu 185
QY 425 CTGCGCTTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCACACCCCTTCGTCCTACGC 484
Db 186 LysAlaAlaGlyGly-----AspPheLysIlePhePheProGlyPheCysLys 201
QY 485 AAATGGCGGCACTCGGCTACTGACCGACTTCCCGCTCAACCGCGGATGCACAAAC 544
Db 202 -----TyrIleAsn-----LeuArgIleAsnTyrArgAsnHisArg 213
QY 545 AAATCCTTTTACGCGCACACCGGCCACCATATCTCGGCGGACGCAATATCGGCGACGAA 604
Db 214 LysIleLeuIleIleAspSerLysValAlaPheLeuGlyGlyPheAsnIleGlyAspGlu 233
QY 605 TACTTCAAGTGGTAGGACACCGCTTTTCGCGGACCTGAGCATCTCGCCACCGGCGACG 664
Db 234 Tyr-----LeuGlyLysAspLys----- 239
QY 665 GTCGTGCGGAAGTATCGCAGCATTCGACCGCTACTGCGCAAGCCATTTCGCCCCACAAAC 724
Db 240 -----AsnIleGlyHis-----Trp-----ArgAspThrHisThr 249
QY 725 GCCAGCGCATATCCGCGCGGCAACATC---GGCAAGGTCTTCAAGCACTCGGATAC 781
Db 250 LysIleLysGlyLeuAlaIleAsnAspLeuGluArgPheLeuLeuAspTrpSerThr 269
QY 782 AACGACGAACATCCAGACACGCG---CTCCTGGCTACCGCGAAACCGTGCAGAACGTCG 838
Db 270 AlaAsnGluSerAspLeuAspIleAspLeuLysTyrPheIleAsnProHisSerThr 289
QY 839 CCCCTCTACCAAAAAATACACGCGGACGTCGACTGCGCAGCGCTCCAAACCCGCTG 898
Db 290 AspLeuProLysLysIle-----IleGlyAlaGlnIle 300
QY 899 ATCAGCGACACCCCTTGCAAAAGGACTCGACCGCGACCGCCCGCAACACCGCGGATTCGCGG 958
```

Percent Similarity: 40.4% Conservative: 56
 Best Local Similarity: 27.3% Mismatches: 162
 Query Match: 9.5% Indels: 93
 DB: 2 Gaps: 18

US-10-665-990A-13 (1-1561) x H82711 (1-467)

Qy	218	CTGCTCGACGACCCGCCAGAGCCCTTTGCCCGCGCGCCCTTATCGAATCTGCCGAA	277
Db	114	LeuLeuAspAspTyrGlnSerSerLeuArgLeuLeuIleAsnAspIleAspSerAlaCys	133
Qy	278	CACAGCCTCGATTGGCAATACTACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTC	337
Db	134	GluArgValHisLeuLeuTyrTyrLeuMetPheAspAspAlaValGlyAspAlaValVal	153
Qy	338	AACCTCATGTACCTTGGCGGAGACGGCGCTGCGGTACGCTGCTGTGGACGAC	394
Db	154	AlaAlaLeuGlnArgAlaAlaAlaArgGlyValSerCysArgLeuLeuLeuAspAlaVal	173
Qy	395	AACAACACGGCGGGTTGGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAA	454
Db	174	GlyAlaLysArgGlyLeuArgAlaTyrGlnGlnCysLeuArgAlaHis--GlyValGlu	192
Qy	455	GTGGCCTGTTCAACCCCTTGCTCTTACGCAAAATGGCGCGCACTCGGGTACTCGTACC	514
Db	193	ValHisAlaMetLeuProGlyGlyLeuArg--TrpArgArgSerGly-----	207
Qy	515	TTCCCCCGCTCAACCGCGCATGCACAAACAAATCCTTTACGCGGCAACCGGCCACC	574
Db	208	-----ArgMetAspLeuArgAsnHisArgLysIleAlaValIleAspAsnArgValAla	225
Qy	575	ATACTCGCGGAGCGCAATATTCGGCGAGCAATCTTCAAGTCGGTGGAGACACCGTTTC	634
Db	226	TyrIleGlySerGlnAsnLeuAlaCysProAspPheVal-----ProHisPhe	241
Qy	635	GCGCATCTGCATCTCTCGCACCGGCGAGCGTCTCGGC---GAAGTATCGCAGCACTTC	691
Db	242	ProAsnArgGluLeuValAla-----ArgValCysGlyProAlaValSerHis-----	257
Qy	692	GACCGTACTGGCGAGCCATTCGCGCCCAACAGCCCGCATCATCTCGGAGCGGCAAC	751
Db	257	-----	257
Qy	752	ATCGGCAAGGCTCTTCAAGCACTCGGATACAAACGAC-----GAAACATCCAGACAC	802
Db	258	-----LeuGluAlaValPheLeuSerAspTrpPheValGluThrGlyGlnArg	273
Qy	803	GCCTCTCTGCTTACCGCGAAACCGTTCGAAACGTCGCTCTTACCAAAAAATACAGACG	862
Db	274	-----LeuArgMetArgSerAspVal-----ProIleCysGlu-----	284
Qy	863	GGAGCATCGATGGCGAGCGCTCCAAACCGGCTGTATCAGGACACCCCTGCCAAAAGGA	922
Db	285	-----GlnAsnValAlaAlaGlnLeuValAlaSerGlyProAlaTyrPro	299
Qy	923	CTCGACCGGACCGCGCAACCCCGGATTGCGGGAGGTGCAAGACCGCTCAAAACAG	982
Db	300	TyrGluAsnAlaArg-----AspAlaValAsnThrMetIleHisLeu	313
Qy	983	CCGCAAAAAGCGTCTATCTGTGTTTCAACCTATTTCCTCCCTACAAAATCCGGCACAC	1047
Db	314	AlaArgArgValValLeuThrThrProTyrPheValProAspGluAlaThrLeuSer	333
Qy	1043	GCACCTGGCAAACTGGTGCAGGAGCGCATAGAGCTT---ACCGTCTCGACCACTCGCTA	1099
Db	334	AlaLeuArgIleAlaGlyThrSerGlyValAspValGlnLeuLeuLeuSerGluSerAsn	353
Qy	1100	CAGGCGACCGCTGCGCGCGTCCATTCCTCGGCTACGTCAAATACCGAAACCGCTGCTC	1159
Db	354	AsnGlnPheLeuMetAlaTrpAlaGlnGluAlaTyrPhe-----GluGluLeuLeu	370
Qy	1160	AAAGCCGGCATCAAACTCTACGAGCTGCAACCAACCATCGCTCCCGCCACAAAAAGAC	1219

Db	354	AsnGlnPheLeuMetAlaTrpAlaGlnGluAlaTyPhe-----GluGluLeuLeu	370
QY	1160	AAAGCGCGCATCAAACTCTACGAGCTGC AACCCCAACCATGCGTCCCGCCCAAAAAGAC	1219

```

Db      371 ArgAlaGlyIleLysIleAlaLeuTyArgProHisPhe----- 383
Qy      1220 AAAGCCTGACGGCAGCTCCGTAAACAGCTGATGCAAAACCTTCATTGTGTGACGGC 1279
Db      384 -----LeuHisAlaLysHisLeuSerValAspAsp 393
Qy      1280 AAACGCATCTTCATCGGCTCATTCAACCTCGACCCCGTTCGGCAGCGCTCAATACCGAA 1339
Db      394 AspIleAlaLeuPheGlySerIleAsnLeuAspIleArgSerPheAlaLeuAsnAlaGlu 413
Qy      1340 ATGGCGGTGCGTCATCGAAAGCCCAAAATCGACAGATGCGACGCGCACCCTCGCCGAT 1399
Db      414 IleGlyMetIleCysTyArgSerArgIleValSerArgLeuGluValGluSerAsp 433
Qy      1400 ACCACACCGCAATACGCTACCGGCTTACCTCGACAAACACACACCGCTCGCATGGCAC 1459
Db      434 -----TyrLeuAlaAsnAlaCysGlnLeuAsnLeuValGluTrpGln 447
Qy      1460 GAT---CCGCGCACCCGAAAAACC 1480
Db      448 GlnArgProAlaTrpArgArgSer 455

RESULT 31
T02345
hypothetical protein KIAA0324 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02345
R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
submitted to the EMBL Data Library, March 1998
A:Description: Sequencing of human chromosome 16p13.3.
A:Reference number: Z14664
A:Accession: T02345
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1791 <RIC>
A:Cross-references: UNIPROT:060382; UNIPARC:UPI00000730AE; EMBL:AC004493; NID:g2996648;
C:Genetics:
A:Map position: 16
A:Introns: 1610/2; 1706/2
A>Note: KIAA0324

Alignment Scores:
Pred. No.:      2,97e-08      Length:      1791
Score:          268.50      Matches:      131
Percent Similarity: 37.6%      Conservative: 51
Best Local Similarity: 27.1%      Mismatches: 217
Query Match:      9.4%      Indels:      85
DB:              2          Gaps:      18

US-10-665-990A-13 (1-1561) x T02345 (1-1791)
Qy      172 GCAATCCGGCACACCCCTCATACAAACGGGTATCCGACATCTACCTGCTCGACGACC 231
Db      689 SerSerProGluHisProProlysSerArgThrAlaArgGlySerArgSerPro 708
Qy      232 ---CCACGAAGCCCTTGGCGC-----CCGCGCCGCCCTTATCGAATCTGCGCAACACAG 282
Db      709 GluProLysThrLysSerArgThrProProArgArg-----ArgSerSerArgSerSer 726
Qy      283 CCTCATTTGCATACTACTATTGGCGCAACGACATTTCCGCGCAGCTGCTGTTCAACCT 342
Db      727 ProGluLeuThrArgLysAlaArgLeuSerArgArgSerArgSerAlaSerSerPro 746
Qy      343 CATGTACCTTGGCGCAGA---ACGCGCGCTGCGCGCTAGCCCTGCTGTTGGACGACAA 399
Db      747 GluThrArgSerArgThrProProArgHisArgArgSerProSerValSerSerProGlu 766
Qy      400 CACGCGCGGTTGGACGATCTCCTGCTCGCCCTCGACAGCCATCCCAATATCGAATGCGG 459
Db      767 ProAlaGluLysSerArgSerArgArgArgArgSerAlaSerSerProArgThr--- 785

```

```

Qy      460 CCTGTTCAACCCCTTCGTCTCTACGCAATGGCGCGCACTCGGCTACCTGACCGACTTCCC 519
Db      786 -----LysThrThrSerArgArgGlyArgSerPro 795
Qy      520 CCGCTCAACCGCGCATGCACAAACAAATCCTTTACCGCCGACCAACCGCGCCACCATACT 579
Db      796 SerProLysProArgGlyLeuGlnArgSerArgSerArg---ArgGluLysThr 814
Qy      580 CGCGCGACGCNAATATCGCGACGAATACTTCAAGTGGTGAGGACACCGGTTTCGCCGA 639
Db      815 ArgThrThrArgArgAspArgSerGlySerGlnSerThrSerArgArgGln 834
Qy      640 CCTGGACATCTCGCCACCGCGCAGCTGTCGGCGGAAGTATCGCAGCACTTCGACCGCTA 699
Db      835 ArgSerArgSerArgSerArgValThrArgArgArg----- 847
Qy      700 CTGGCGAAGCATTCCTCCGCCCAACCGCATCATCCGACGCGCAACATCGGCA 759
Db      848 -----GlyGlySerGlyThrHisSerArgSerProAlaArgGln 860
Qy      760 GGGTCTTCAACGACTCGGATACAAAGAGAAACATCCAGACCGGCTCCT-----CGG 813
Db      861 GluSerSerArgThrSerSerArgArgGlyArgSerArgThrProProThrSer 880
Qy      814 CTACCGGAAACCGCTCGAACAGTCCGCCCTCTACCA-----AAAAAT 855
Db      881 ArgLysArgSerArgSerArgThrSerProAlaProTrpLysArgSerArgSerArgAla 900
Qy      856 ACAGCGGGAGCATCGATCGATGGCAGCGTCCAAACCGCCTGATCAGCAGCACCCCTGC 915
Db      901 SerProAlaThrHisArgSerArgSerArgThrProLeuIleSerArg----- 917
Qy      916 AAAGAGACTCGACCGCGACCGCCGCAACCGCATTCGCGGAGGTGCAAGACCGCT 975
Db      918 ArgArgSerArgSerArgThrSerProValSerArgArgArgSerArgSerArgThrSer 937
Qy      976 CAACAGCCCGCAAAAGCGTCTATCTGTTTCACCCCTATTTCGTCCCTCAAAATCCGG 1035
Db      938 ValThrArgArgSerArgSerArgAlaSerProVal-----SerArg 952
Qy      1036 CACAGCGCACTGGCAAAACTGGTGCGAGCGGATAGACCTTACCTGCTGACCAATC 1095
Db      953 ArgArgSerArgSerArgThrProProValThrArgArgArgSerArgSerArgThrPro 972
Qy      1096 GCTACAGCGCAGCAGCTTGGCGCGTCCATTCGGCTACCTCAAAATACCGAAACCGCT 1155
Db      973 ThrThr-----ArgArgSerArgSerArgThrProProValThrArg 987
Qy      1156 GCTCAAAGCGGCATCAAACCT-----CTACGAGCTGCAACCCAA 1194
Db      988 ArgArgSerArgSerArgThrProProValThrArgArgArgSerArgSerArgThrSer 1007
Qy      1195 CCATGCCCTCCCGCCACAAAAGAGAGGCTGACCGGAGCTCGTAAACAGGCTGCA 1254
Db      1008 ProlIleThrArgArgSerArgSer-----ArgThrSerProVal 1021
Qy      1255 TGCMAAACCTTCATTGTGGACGCAACAGCATCTTCATCGCTCATTCACCTCGACCC 1314
Db      1022 ThrArgArgSerArgSerArgThrSerProValThrArg----- 1035
Qy      1315 CCGTTCGCACGGCTCAATACCGAAATGGGCGTCTCATCGAAAGCCCAAAATCGCAGA 1374
Db      1036 -----ArgArgSerArgSerArgThrSerProValThrArgArgArgSerArgSerArg 1053
Qy      1375 ACAGATGAGCGCACCCCTCGCGGATACCAACCGGAATACGCTTACCGGTTACCTCGA 1434
Db      1054 ThrProProAlaIleArgArg-----SerArgSerArgThrProLeuLeuProArg 1071
Qy      1435 CAACACAAACCGCTGCAATGCGCATCCCGCCACCGCAAAACCTACCCGCAAGAAC 1494
Db      1072 Lys-----ArgSerArgSerArgSerProLeuAlaIleArgArgSer 1086
Qy      1495 CGAAGCCAAACTTTGGAAACGCATCGCGCGCAAAATCTATC-----CCTGCT 1542

```


Db 1087 ArgSerArgThrProArgThrAlaArgGlyLysArgSerLeuThrArgSerProProAla 1106
Qy 1543 GCCATCGAAGG 1554
Db 1107 IleArgArgArg 1110

RESULT 32

B82971
cardiolipin synthase PA5394 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B82971
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:2043737; PMID:10984043
A:Accession: B82971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <STO>
A:Cross-references: UNIPROT:Q9HTH0; UNIPARC:UPI000000C5FDF; GB:AB004952; GB:AB004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: cls; PA5394
C:Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.: 2,88e-08 Length: 490
Score: 268.00 Matches: 111
Percent Similarity: 37.7% Conservative: 60
Best Local Similarity: 24.5% Mismatches: 186
Query Match: 9.4% Indels: 96
DB: 2 Gaps: 13

US-10-665-990A-13 (1-1561) x B82971 (1-490)

Qy 95 TGGTTGCCCCCAGGAGGACGCGGAGGCGCTATTTCATATCTCCAACTGTC 154
Db 97 TrpArgProTrpValGluGluAlaLeuAlaArgValSerGlyTyrLysGly--- 115
Qy 155 CTCCTGGACACATCTCTGCAATCCGCGCACACCCCT-----CATAACAACGGGCTATCC 208
Db 116 ---LeuLysAlaLeuValArgMetThrArgThrProThrLeuAlaAsn--- 131
Qy 209 GACATCTACCTGCTGAGGACCCCGACGAGCCCTTCCGCGCGCGCCCTTATCGAA 268
Db 132 ArgValArgLeuLeuValAsnGlyGluAlaSerPheGluAlaMetPheLysAlaIleSer 151
Qy 269 TCTGCCGAACACAGCCCTCGATTGTCATATCTACATTTGGCGCAACGACATTTCCGCGCAGG 328
Db 152 AlaAlaArgGlnValIleLeuValGlnPhePheIleValArgAspAlaLeuGlyGln 171
Qy 329 CTGCTGTTCACCTCATGTACTTTCGCGCAGACGCGCGTGGCGGTACGCTGCTGTTG 388
Db 172 ArgLeuGlnGlnLeuLeuLeuGluArgAlaAlaAsnGlyValGluValPhePheLeuTyr 191
Qy 389 GACGACAAACACAGCCCGGGTGGACGATCTCTGCTAGCAATGGCGGCGGACATCCCAAT 448
Db 192 AspAlaIleGlySerHisAlaLeuProHisArgTyrValGluArgLeuArgGlnGly 211
Qy 449 ATCGAAGTCGCGCTGTTCAACCCCTTCGCTAGCAATGGCGGCGGACATTCGCTACCTG 508
Db 212 ValGlnMetHisGlyPheSer-----ThrGlySerGlyMetLeu 224
Qy 509 ACCGACTTCCCGCCCTCAACCGCGGATGACAAACAAATCTTTACCGCGGCAACCGC 568
Db 225 AsnArgPhe---GlnValAsnPheArgAsnHisArgLysValValValValAspGlyGlu 243
Qy 569 GCCACCATATCTCGCGGACGCAATATCGGCGACGAATACTTCAAGTCGGTGAGGACACC 628
:::||||| |||::| ||||| |||||

Db 244 CysGlyPheValGlyGlyHisAsnValGlyValGluTyr-----LeuGlyGluLysPro 261
Qy 629 GTTTTCGCC-----GACCTGGACATCTCCGCCCGGACGCGTCTGGCGAAATA 679
Db 262 ProLeuAlaProTrpArgPheMetGluLeuArgGlyProAlaValAlaCysLeu 281
Qy 680 TCGCACGACTTC-----GACCGCTACTGGCAAGCCATTCCGCCCAACACGCCACGCGC 733
Db 282 GlnGluSerPheAlaGluAspTrpTyrTrpAlaThrHisSerLeu----- 296
Qy 734 ATCATCCGCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACACGACGAAACA 793
Db 296 ----- 296
Qy 794 TCCAGACACGCGCTCTCGCTACCGGAAACCGTCGCAACAGTCGCGCCCTTACCAAAA 853
Db 297 -----ProProLeuIleLeuProGlnTyrAspSerGluGlyAlaLeuCysGlnVal 314
Qy 854 ATACAGACGGGACGATCGACTGGCAGCGGTCCAAACCCCGCTGATCAGCGACACCCCT 913
Db 315 ValAlaSerGlyProAlaAspAlaGlnGluThrCysSerLeuPhe----- 329
Qy 914 GCAAAAGACTCGACCGCGACCGCGCAACCGCGGATTCGCGGAGGCTCAGACGCG 973
Db 330 -----PheValGluMet 333
Qy 974 CTCAAACAGCCGAAAGCGTCTATCTGGTTTCACCCCTATTTCTGCTCCCTACAAATCC 1033
Db 334 IleAsnAlaAlaHisGluArgValTrpIleThrSerProTyrPheValProAspGluAla 353
Qy 1034 GGCACAGCAGCTTGGCAAACTGTCGAGGAGGATAGACGTTACGTTCTCTGACCAAC 1093
Db 354 ValMetAlaAlaLeuArgLeuAlaValLeuArgGlyValAspValArgLeuLeuIlePro 373
Qy 1094 TCGCTACAGCGACCGAGCTTGGCGCGTCTCATTTCCGGCTACGTCAAATACCGAAAAACCG 1153
Db 374 Ser-----ArgProAspHisArgThrValTyrAlaAlaSerSerLeuTyrAlaLeuGlu 391
Qy 1154 CTGCTCAACCGCGCATCAACTCTACGAGCTGCACCCCAACCATCGCGTCCCGCCACA 1213
Db 392 AlaIleArgAlaGlyValLysValPheArgTyrGlnProGlyPhe----- 406
Qy 1214 AAAGACAAAGCCCTGACCGCGAGCTCCGTAAACAGCGCTGCATGCCAAACCTTCATTGTG 1273
Db 407 -----LeuHisGlnLysValValLeuVal 414
Qy 1274 GACGGCAACGCAATCTTCATCGGCTCATTCAGCTCGACCCCGCTTCGCGACGCTCAAT 1333
Db 415 AspArgAspThrAlaAlaValGlySerAlaAsnLeuAspAsnArgSerPheArgLeuAsn 434
Qy 1334 ACCGAATGGCTGCTCATCGAAAGCCCAAAATCCGACAGACAGATGGAGCGCACCCCTC 1393
Db 435 PheGluValMetValThrValAspGluGlyPheAlaGlyGluValGluAlaMetLeu 454
Qy 1394 -----GCCGATACCAACCCCGAA 1411
Db 455 GluAlaAspPheAlaGluSerLeuGluPheThrProGlu 467

RESULT 33

S60089
cardiolipin synthetase homolog ywie - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S60089; S5414; F70059
R:Crux Ramos, H.; Boursier, L.; Moszer, I.; Kunst, F.; Danchin, A.; Glaser, P.
EMBO J. 14, 5984-5994, 1995
A:Title: Anaerobic transcription activation in Bacillus subtilis: identification of dis
A:Reference number: S60080; MUID:96112813; PMID:8846791
A:Accession: S60089
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-500 <CRU>
A:Cross-references: UNIPROT:P45860; UNIPARC:UPI0000060BCF; EMBL:Z49884; NID:G971335; PI

A>Note: host Equus caballus (domestic horse)
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C/Accession: H36802
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank March 1992

A>Description: The DNA sequence of equine herpesvirus-1.

A/Reference number: A3605

A/Accession: H36802

A/Molecule type: DNA

A/Residues: 1-797 <TEL>

A/Cross-references: UNIPROT:P28968; UNIPARC:UPI0000138750; GB:M86664; MID:g330791; PIDN:

R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

Virology 189, 304-316, 1992

A/Title: The DNA sequence of equine herpesvirus-1.

A/Reference number: A41831; MUID:92295566; PMID:1318606

A/Contents: annotation; possible protein-coding frames

A>Note: neither amino acid nor nucleotide sequence is given

C/Genetics:

A/Gene: 71

C/Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein hom

C/Keywords: glycoprotein; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-797/Product: glycoprotein X #status predicted <MAT>

F:23-465/Region: serine/threonine-rich

F:489-797/Domain: equine herpesvirus 1 glycoprotein homology <EHG>

F:766-790/Domain: transmembrane #status predicted <TMN>

F:590/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	3.92e-08	Length:	797
Score:	266.00	Matches:	139
Percent Similarity:	36.2%	Conservative:	52
Best Local Similarity:	26.3%	Mismatches:	206
Query Match:	9.3%	Indels:	131
DB:	1	Gaps:	20

US-10-665-990A-13 (1-1561) x VGBEX1 (1-797)

Qy	20	TCTGAAACTATATCCCGATGAAACACGACGAGCCTCATTTCCCTTTTATGCTCCTCTC	79
Db	52	SerProThrThrSerProProThrThrSerSer-----	62
Qy	80	TGTTTCATGTTCTTCATGGTTCCTCCCTGACGAGACGAGACGAGACGAGCCGTCATTTCAAT	139
Db	63	-----SerProProThrThrSerThrHisThrSerSerProSer---Ser	75
Qy	140	ACTTCCAAACCTGTCTCTCTGGACACATCTCTGCAATTCGGACACCCCTCATATA	194
Db	76	ThrSer-ThrGlnSerSerThrAlaAlaThrSerSerSerAlaProSerThrAlaSe	95
Qy	195	----ACACGGGCTATCCGACATCTACTGCTGACGACGCCACGACGAGCCCTTGCGGCC	250
Db	95	rSerThrThrSerIleProThrSerThr-----SerThrGluThrThrThrThrProTh	114
Qy	251	CGCGCCGCCCTTATCGAATTCGCGACACAGCCTCGATTTCGAATTAATTAATTTGGCGC	310
Db	114	rAlaSer-----ThrThrThrProThrThrThrAlaAlaProThrThrAlaAlaTh	132
Qy	311	AACGACATTCGCGCAGGCTGCTTCAACCTCATGTACCTTCCGCGAGAACGGGGGTG	370
Db	132	rThrThrAlaValThrThrAlaAlaSerThrSer-----	143
Qy	371	CGCGTACGCTGCTGTTGGACGACAAACACGCGCGGGTGGACGATCTCTGCTGCC	430
Db	144	-----AlaGluThrThrThrThrAlaThrAlaThrAlaThrAlaThrSerThrPr	157
Qy	431	CTGACAGCATCCCAATATCGAAGTCGCGCTGTTCACCCCTTTCGTCCTACGCAATGG	490
Db	157	oThrThrThrThrProThrSerThrThrThrThrThrAlaThrThrThrValProThrTh	177
Qy	491	CGGCACTCGGCTACCTGACGAGTCTCCCGGCTCAACCGCGCATGACACAAATCC	550
Db	177	rAlaSerThrThrThrAspThrThrThrAlaAlaThrThrThrAlaAlaThrThrAl	197

Qy	551	TTTACCCTGACACCGCGCACCATACTCGCGGAGCAATATCGGAGCAATCTTC	610
Db	197	aAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaThrThrAl	217
Qy	611	AAAGTCGGTGAGGACACCGTTCCTCGCGACCTCGGACATCTCTCGGACCGCGAGGCTGCTC	670
Db	217	aAla-----ThrThrThrAlaAlaThrThrSerSe	227
Qy	671	GGCGAAGTATCGCAGCAGCTTCGACCGCTACTGCGCAAGCCATTCGCGCCACCAAGCCACG	730
Db	227	rAlaThrThrAlaAlaThrThrThrAlaAla-----ThrThrThrAl	241
Qy	731	CGCATCATCGCAGCGCAACATCGCAAGGCTCTCAAGCACCTCGGATACCAAGCAGCAA	790
Db	241	aAlaThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrGl	261
Qy	791	ACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCGAACAGTCGCGCCCTCTACCAA	850
Db	261	ySerProThrSerGlySerThrSerThrThrGlyAlaSerThrSerThrProSer-----	279
Qy	851	AAAATACAGAGGAGCCATCGACTGCAGAGCGTCCAAACCCCTGTATCAGGACAC	910
Db	280	-----AlaSerThr-----AlaThrSerAlaThrPr	288
Qy	911	CCTGCAAAAGGACTCGACCGCGCGCAACCGCGATTCGCGGAGGCTGCAAGAC	970
Db	288	oThrSerThrSerThrSerAlaAlaThr-----	298
Qy	971	CGCGTCAACACGCGGAAAGCGTCTATCTGTTTTCACCCCTATTCGTCCTACAAA	1030
Db	299	---SerThrProThrProThrSerAlaAlaThrSerAlaGluSerThrThrGluAl	317
Qy	1031	TCGCGCACACGACTGCGCAAAACTGGTGACGAGCGCATAGACGTTACGCTCTGACC	1090
Db	317	aProThrSerThr-----Pr	322
Qy	1091	AACTCGCTACAGCGCGACCGCTTTCGCGCTCCATTCGCGCTACGTCACAAATACGAAA	1150
Db	322	oThr-----ThrAspThrThrThrProSerGluAlaThrThrAlaThrThrSerPr	339
Qy	1151	CCGCTGCTCAAGCCCGCATCAAACTCTACGAGCTGCAACCCCAACCATG-----	1199
Db	339	oGluSerThrThrValSerAlaSerThrThrSerAlaThrThrThrAlaPheThrGl	359
Qy	1200	-----CCGTCGCGCCAAAGAGCAAGGCTGACCGGAGCTCCGTAACCGAGCTG	1252
Db	359	uSerHisThrSerProAspSerThr-----	368
Qy	1253	CATGCCAAACCTTCATTTGGAGCGCAACCGCATCTTCATCG-----GCTCATTC	1303
Db	369	-----GlySerThrThrThrAlaGluProSerSerThrPheThrLeuThrProSe	385
Qy	1304	AACCTGACCCCGTTCGCGACGCTCAATACCGAAATGGCGCTGTCATCGAAAGCCCC	1363
Db	385	rThrAlaThrPro-----SerThrAspGlnPheThrGlySerSerAlaSerTh	401
Qy	1364	AAATCGCAGACAGATGAGCGCACCTCGCGATACACACCCGAATACGCTACCGC	1423
Db	401	rGluSerAspSerThrAspSer-----SerThrValProThrThrGlyThrGluSerI	419
Qy	1424	GTTACCTCGCAAAACACACCGCCTGC-----AATGGCAGTCCCGCC	1468
Db	419	eThrGluSerSerThrThrThrGluAlaSerThrAsnLeuGlySerSerThrThrTy	439
Qy	1469	ACCCGAAAAACCTACCCGACGAAACCGAGCCAAACTTTGGAACGATCGCGCGAAA	1528
Db	439	rThrGluAlaLeuGluThrProAspGlyAsnThrThrSerGlyAsnThrThrProSerPr	459
Qy	1529	ATCCTATCCCTGCTGCCATCG	1550
Db	459	oSer---ProArgThrProSer	465

RESULT 35
AE3539
cardiolipin synthetase (EC 2.7.8.-) [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AE3539
R;DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <KUR>
A;Cross-references: UNIPROT:Q8YD8; UNIPARC:UPI00000583DD; GB:AE008918; PIDN:AAL53480.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10239
A;Map position: II
C;Superfamily: cardiolipin synthase
C;Keywords: transferase

Alignment Scores:
Pred. No.: 5,3e-08 Length: 492
Score: 263.50 Matches: 98
Percent Similarity: 39.2% Conservatives: 69
Best Local Similarity: 23.0% Mismatches: 180
Query Match: 9.2% Indels: 79
Gaps: 13

US-10-665-990A-13 (1-1561) x AE3539 (1-492)

```
QY 128 GTCATATTCAATACATCCAAACCTCTCTCTGGACAACTCTGCAAAATCGGCACACC 187
Db 106 ArgGlnPheGlyAlaMetLys---lleLeuGlyAspAlaValSerLeuTyAspPheThr 124
QY 188 COTCATAACAGGGGTATTCGACATCTACCTGCTCGACGACCCCAAGAGCCCTTGGC 247
Db 125 SerGlyAsnThr-----lleGluMetLeuGluAlaGlyAspGluAlaTyrAla 140
QY 248 GCCCGCGCGCCCTTATCGAATCTCCGACACACACGCTCGATTCGCAATACATACATTTGG 307
Db 141 AlaMetLeuGlyAlaileGlyArgAlaGluArgSerileValLeuGluThrTyrilePhe 160
QY 308 CGCAACGACATTCGCGGAGGTGCTGTTCAACCTCATGCTACCTTCCCGCAGAACGCGGC 367
Db 161 AspHisAspAlaileGlyLysLysPheAlaAspAlaLeuGlyAspAlaValGlnArgGly 180
QY 368 GTGCGCGTACGCTGCTGTGGACGACAAACACGCGGGGTGGACGATCTCTGCTC 427
Db 181 ValGluValArgValLeuValAspAlaValGlyAlaArgTyrSerPheProSerileVal 200
QY 428 GCCCTCGACAGCCATCCCAATATCAAGTGCCTGTTCAACCCCTTCTGCTCTACGCAA 487
Db 201 LysLeuLeuLysGluLysGlyValLysValAlaValPheAsnGlyAsnlelleleGly 220
QY 488 TGGCGCGCCTACCTGCTACCGACATTCCTCCCGCCTCAACCGCGCATGCAACAA 547
Db 221 LeuArg-----LeuProTyrAlaAsnLeuArgThrHisArgLys 233
QY 548 TCCTTTACGCGCAACACGCGCCACCATATCTCGGCGGACGCAATATCGGCGCAATAC 607
Db 234 MetLeuileValAspGlyLysThrAlaPheThrGlyGlyMetAsnlelleArgAlaGlyPhe 253
QY 608 TTCAAA---GTGCGGTGAGGACACCGCTTTTCGCGACCTGGACATCTCTCGCCACCGGACG 664
Db 254 ValArgAlaileAlaGlyAspAlaValAlaPheAspThrHisPheLysLeuGluGlyPro 273
QY 665 GTGTCGCGGGAAGTATCGCAGCATTCGACCGCTACTCGGCAAGCCATTCGCCCCCAAC 724
Db 274 AlaileAlaAspLeuPheHisileAlaSerGluAspTyrArg----- 287
```

```
QY 725 GCCAGCGCATCATCCGACGCGGCAACATCGGCAAGGTTCTTCAAGCACTCGGATACAC 784
Db 288 -----PheAlaThrGlyGluLeuThrGly----- 296
QY 785 GACGAAACATCCAGACAGCGCTCTCGGTACCGCGAAACCGTCGAACAGTCGCGCCCTC 844
Db 297 -----GluAlaTyrSerileAlaProPro 304
QY 845 TACCAAAAAATACAGACGGGACGATCGATCGTGGCAGAGCGTCCAAACCCGCTGATAGC 904
Db 305 GluAsnProProGlyThrGlyThrLeu-----ValArgValValGly 318
QY 905 GACACCCCTGCAAAAGGACTCGACGCGGACCGCGCAACCGCGGATTCGCGGAGGTG 964
Db 319 SerGlyProAspLysAsnLeuGluThrAsnHisArg-----MetMet 332
QY 965 CAAGACGCGCTCAAAACAGCCGCAAAAGCGCTCTCTGTTTTCACCTATTTCTGCTCCT 1024
Db 333 MetGlyAlaPheSerileAlaGlnHislellelleMetThrProTyrLeuLeuPro 352
QY 1025 ACAAATCCGCGACAGACGCTGCGCAAACTGTGTGAGGACGATAGAGCTTACCGTC 1084
Db 353 AspArgGluLeuSerAlaLeuValThrAlaAlaArgArgGlyValSerValAspIle 372
QY 1085 CTG-----ACCACTCGCTACAGCGACCGCGCTGCGCGCTCCATTTCGCGTAC 1135
Db 373 ValValProGlyValAsnAsnLeuLysLeuValAsp-----ArgAlaMetArgAlaGlnPhe 391
QY 1136 GTCAATATACCGAAACCGCTCTCAAAAGCGCGCATCAAACTCTACGAGCTGCAACCCAAC 1195
Db 392 -----AspGlnLeuLeuArgAspGlyCysArgIleTyrArg----- 403
QY 1196 CATCGCTCCCGCCCAACAAAGCGCTGACGCGCAGCTCGTAACAGCCCTGCAT 1255
Db 404 -----AlaGlyGlyAlaPheAsn-----His 410
QY 1256 GCCAAACCTTCATTGTGAGCGCAACCGCATCTTCATCGGCTCATCAACCTCGACCCC 1315
Db 411 SerLysLeuMetThrileAspGlyAlaTyrSerTyrValGlySerSerAsnleleAspPro 430
QY 1316 CGTTCGCGCAGCGCTCAATACCGAAATGGCGCTCGTTCGTAAGGCCCAAAATTCGAGAA 1375
Db 431 ArgSerLeuArgLeuAsnPheGluValAspLeuGluLeuLeuAspArgAspValAlaArg 450
QY 1376 CAGATGAGCGCACCTC 1393
Db 451 GlnValGluGluArgIle 456
RESULT 36
E97307
probable cardiolipin synthase (phospholipase D family) [imported] - Clostridium acetobu
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97307
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-510 <KUR>
A;Cross-references: UNIPROT:Q97B04; UNIPARC:UPI000000756C; GB:AE001437; PIDN:AAK81248.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3316
C;Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 6.08e-08 Length: 510
Score: 262.50 Matches: 119
Percent Similarity: 36.5% Conservatives: 82
```

Best Local Similarity: 21.6% Mismatches: 211
 Query Match: 9.2% Indels: 138
 DB: 2 Gaps: 19

US-10-665-990A-13 (1-1561) x E97307 (1-510)

```

Qy 28 TATATCCCGATGAAAAACACGAGCCTCAT----- 57
Db 54 TyrIleLeuTyrAsnAsnSerAsnProAlaTyrLysIleValTrpIleProIleLeu 73
Qy 58 ---TTCCCTTTTATGCTCTCTCTGTTTCATGTTCTTCATGTTGTTGCTGCTGGAAGA 114
Db 74 ValPheProLeuPheGlyLeuPheGlyLeuPheGlySer-LysLeuSerAr 93
Qy 115 ACGACGGAAGCCGCTCAT-----TTCAATACTTCCAAACCTGCTCCTCGGACAA 165
Db 93 gArgThrArgLysLysMetGluCysIleValAlaLysSerArgAlaValLysGlu 113
Qy 166 CATCTGCAATTCGGGCACACCCCTCATACACGGGCTATCC----- 208
Db 113 yGluValIleAsnGluIleArgAsnAsnGluThrAlaAlaAsnGlnSerSerTy 133
Qy 209 ---GACATCATCTGCTCGAGACCCCGACGAGCCCTTCCGCCCGC----- 253
Db 133 rIleGlnAsnTyrSerLeuPheProProTyrLysTyrThrLysAlaGluTyrPheSerTh 153
Qy 254 -----GCCGCCCTTATCGAA-----TCTGCCGAACACAGCCTCGA 288
Db 153 rGlyGluGluThrPheAsnGluLeuIleGluGlnLeuLysLysAlaLysGlnTyrIlePh 173
Qy 289 TTTGCAATACATTTGGCCACACGACATTCCTCGGAGGCTGCTGTTCAACCTCATGTA 348
Db 173 eLeuGluTyrPheIleIleLysGluGlyValMetTrpAsnSerValLeuGlnIleLeuAr 193
Qy 349 CCTTCCGCAGAACCGCGCTGCGGTACGCTGCTGTGTGACACACACACGCGGG 408
Db 193 gAsPlysValGlnGluGlyValGluValArgValIleTyrAspAsp----- 208
Qy 409 GTTGGAGGATCTCTGCTGCGCCTCGACGAGCATCCCAAT----- 448
Db 209 -----IleGlyCysMetPheThrLeuProAsnGlyTyrHisLysGluLeuGl 224
Qy 449 -----ATCGAAGTGCCTGTTCAACCCCTTCTGCTCATCGCAATGGCGCGCACT 498
Db 224 uLysIleGlyIleLysCysCysValPheAsnProLeuIle----- 237
Qy 499 CGGCTACTGACCGACTTCCCGCTCAACCGCGCATGCACACAAATCCTTTACGCG 558
Db 238 ----ProMetValSerPhe-----LysPheAsnAsnArgAspHisArgLysIleAlaValI 255
Qy 559 CGAACACCGCGCCACCATCTCGCGGACGCAATATCGCGGACGAATATCTTC-----AAAGT 615
Db 255 eAspGlyLeuValGlyPheThrGlyGlyIleAsnLeuSerAspGluTyrIleAsnLysTy 275
Qy 616 CGGTGAGACACCTTTTCGCCGACCTGGACATCCTCGCACCGCGGACGCTGCGCGCA 675
Db 275 rGlnLysTyrGlyTyrTrpLysAspThrAlaIleLysValGluGlyLysAlaAlaTrpAs 295
Qy 676 AGTATCGACGACTTCCACGCTACTGGCAAGCCATTCGCCCACACGCGCAGCCGAT 735
Db 295 nLeuSerValMetPheLeuSerMetTrpSer-----Ph 306
Qy 736 CATCCGCGCGGCAACATCGCAAGGTCTTCAAGCACTCGGATACACACGACGAACATC 795
Db 306 eLeuArg-----GlyIleAspGluAspPheAsnThrPheLysLysAsnIleGluLysGl 324
Qy 796 CAGACACCGCTCTGCGCTACCGCGAAACCGTCGAACAGTCGCCCTCTTACCAAAAAAT 855
Db 324 uPheSerSerArgGluGlyTyrValGlnProPheAlaAspSerProLeu----- 340
Qy 856 ACAGACGGGACGATCATCGTGGACAGCGGTCCAAACCGGCTGATCAGCGACACCCCTGC 915
Db 341 ----AspGlyGluProValGlyGluIleIleTyrMetAsnLeuIleSer----- 355

```

```

Qy 916 AAAAGGACTCGACCGCGACCGCGCAAAACCGCGATTGCGGGAGGCTGCAAGACGCGCT 975
Db 355 ----- 355
Qy 976 CAAAACAGCCGAAAAAGCGTCTATCTGTTTTCACCCCTATTTCGTCCCTACAAAATCCGG 1035
Db 356 ---LysAlaThrLysTyrValTyrIleThrThrProTyrLeuValIleGlyAsnGluMe 374
Qy 1036 CACAGACGCACTGCGCAAAACTGGTGCAGGACGGCATAGACGTTACCGTCTGACCAATC 1095
Db 374 tValThrAlaLeuThrSerAlaAlaLysGlyValAspValArgIleIleThrProHi 394
Qy 1096 GCTACGCGGACCGACCGTTCGCGCTCATTTCCGGCTACGCTCAAAATACCGAAACCCCT 1155
Db 394 sIle-----ProAspLysLysIleValHisSerValThrLysSerTyrTyrLysValIle 412
Qy 1156 GCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCAACCATCGCGTCCCGCCACAAA 1215
Db 412 uIleGluSerGlyValLysIleTyrGluTyrMetProGlyPhe----- 426
Qy 1216 AGACAAAGGCTGACCGCGACCTCGGTAAACGAGCTGCATCGCAACCAACCTTCATTTGTGGA 1275
Db 427 -----IleHisSerLysThrTyrValCysAs 435
Qy 1276 CGGCAAAAGCATCTTCATCGGCTCATTCACATCGACCCCGTTCGCGACGCGTCAATAC 1335
Db 435 pAsnGluTyrGlyValValGlySerIleAsnMetAspPheArgSerLeuTyrLeuHisPh 455
Qy 1336 CGAAATCGGCTGCTCATCGAAAGCCCAAAATCGACAGACAGATGAGCGCACCCCTCGC 1395
Db 455 eGluCysGlyValTrpMetTyrLysThrAsnThrValTyrAspIleLysLysAspPheMe 475
Qy 1396 CGATACACACACCGAATACGCTTACCGGTACCGTTCACCTCGACAAACACACACCGCTGCAATG 1455
Db 475 tAspThrLeu---AspLysSerLysGluIleThrLeuGluGluIleAsnLysValLysTr 494
Qy 1456 GCACGATCCCGCGCACCGCAAAACCTACCGCAACGACCGCAAGCAACCTTGGAAACG 1515
Db 494 pTyr-----SerThrLeuTrp---Ar 500
Qy 1516 CATCGCGCAAAATCTATCCCTGCTG 1543
Db 500 gValValLeuArgValPheAlaProLeu 509

```

RESULT 37

D83103
 probable phospholipase PA4339 [imported] - Pseudomonas aeruginosa (strain PA01)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C;Accession: D83103
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Lior
 adman, S.; Yuan, Y.; Brody, L.D.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Liu
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Residues: 1-359 <STO>
 A;Cross-references: UNIPROT:Q9HW62; UNIPARC:UPI00000C5CCA; GB:AE004850; GB:AE004091; N
 C;Genetics:
 C;Gene: PA4339
 C;Superfamily: cardiolipin synthase

Alignment Scores:
 Pred. No.: 7,25e-08 Length: 359
 Score: 261.00 Matches: 93
 Percent Similarity: 38.5% Conservative: 57
 Best Local Similarity: 23.8% Mismatches: 148
 Query Match: 9.2% Indels: 92

```
DB: 2 Gaps: 10
US-10-665-990A-13 (1-1561) x D83103 (1-359)
QY 263 ATCGAATCTGCGGAACACAGCCTCATTTGCAATACATTTGGCGCAACGACATTTCC 322
DB 5 IleGluAlaAlaGlnArgSerIleGluLeuGluLeuTyrLeuValGluAspGlyHisCys 24
QY 323 GCGAGGTGCTGTTCAACCTCATGTACCTTCCGCGAAGCGCGGTGCGGTACGCGCTG 382
DB 25 AlaGluLeuPheLeuValAlaLeuLeuAspAlaArgArgGlyValAlaValArgCys 44
QY 383 CTGTTGGACGACAACAACGCGGGGTGGACGATCTCTGCTCGCCCTCGACAGCCAT 442
DB 45 LeuPheAspGlyPheGlyCysLeuGlyLeuGlySerAlaTrpIleGlnArgLeuArgGlu 64
QY 443 CCCAATATCGAAGTCGCGCTGTTCAACCCCTTCGTCCTACGCAATGCGCGCACTCGGC 502
DB 65 AlaGlyGlyGluLeuArgLeuTyrAsnProLeu-----ArgTrpLysLeuThrGly 81
QY 503 -----TACCTGACCGACTTCCCGCGCTCAACCGCGCATGCACAAACAATTCCTTT 553
DB 82 GlyAsnLeuTyr-----ArgAspHisArgLysLeuLeu 92
QY 584 ACCGCGGACAACCGCGCCACCATCTCGCGGAGCGCAATATCGGCGAGCAATACTTCAA 613
DB 93 LeuValAspGlyArgLeuGlyTyrValGlyAlaGlyIleThrAspGluPheTrpGlu 112
QY 614 --GTCCGTGAGGACACCGTTTTCCCGCAGCTGGACATCTCGCCACCGCAGCGTCTGC 670
DB 113 ProValSerAspValSerAlaTrpArgGluValMetValGluMetAspGlyProValVal 132
QY 671 GCGCAAGTATCGCAGCTTCGACCGCTACTGGCAAGCCATTCGCGCCCAACAGCC--- 727
DB 133 AlaAspTrpAlaLeuPheGluArgGlnTrpLeuAlaCysLeuGluGluLysAlaTrp 152
QY 728 -----ACGCGCATCTCCGCGGCGCAACATCGGCAAGGT 763
DB 153 LysProArgGluGlyMetThrLeuThrArgLeuProGlnProGlnProGlyAlaAlaArgGly 172
QY 764 TTTCAGCAGCTCGGATCAACAGCAACATCCAGACGCGCTCTCGCTACCGCAA 823
DB 173 LeuGlyArgValAlaTyrAlaAsp----- 180
QY 824 ACCGTGCAACAGTCGCCCTCTTACCAAAAAATACAGACGGAGCATCGACTGGCAGAGC 883
DB 181 -----Ala 181
QY 884 GTCCAAACCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGCGCGCAAA 943
DB 182 ArgGlnHisArgAspIleLeuGlnSerLeuValArgAlaLeuAsnGlySerArgArgArg 201
QY 944 CGCGCGATTGCGGGAGGCTCAAGACGCGCTCAACACGCGCTCAACACGCGGCTATCTG 1003
DB 202 -----IleTrpLeu 204
QY 1004 GTTTCACCTATTTGCTCCCTACAAAAATCCGCGACAGACGCACTGCGCAAACTGTGCGAG 1063
DB 205 AlaThrProTyrPheLeuProThrTrpLysValArgAlaLeuArgLysAlaAlaGln 224
QY 1064 GACGCGATAGAGTTACGTCGTGACCAACTCGCTACAGGCGACCGAGTTCGCCCGCTC 1123
DB 225 ArgGlyValGluValArgLeuLeuLeuAlaGly---ArgLeuThrAspHisAlaProVal 243
QY 1124 CATTCGCGGTACGTCAAATACCGAAAAACCGCTGCTCAAAAGCGCGCATCAAACTCTACGAG 1183
DB 244 ArgTyrAlaGlyGlnArgTyrTrpArgLeuLeuArgAlaGlyValArgIleHisGlu 263
QY 1184 GTGCAACCCCAACCATGCGCGTCCCGCCACAAGAAAGCAAGGCTGACCGGAGCTCCGTA 1243
DB 264 TyrGlnProArgPhe----- 268
QY 1244 ACCAGCCTGCATGCCAAAAACCTTCATTGTGGAGCGCAACGCATCTTCATCGGCTCATTC 1303
```

```
DB 269 -----LeuHisLeuLysMetValMetValAsp---AspTrpValSerValGlySerCys 285
QY 1304 AACCTCAGCCCCCTTCGCGACGGCTCAATACCGAAATGGCGTCGTCATCGAAAGCCCC 1363
DB 286 AsnPheAspHisTrpAsnLeuArgPheAsnLeuAspAlaAsnLeuGluAlaLeuAspPro 305
QY 1364 AAAATCGCAGAACAGATGGAGCGCACCCCTC 1393
DB 306 AspPheThrAsnGluAlaAlaAlaSerLeu 315
RESULT 38
AH2333
cardiolipin synthase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2333
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2333
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-480 <KUR>
A:Cross-references: UNIPROT:Q8YPH1; UNIPARC:UP100000CEA63; GB:BA0000019; PIDN:BA075922.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4223
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 1.45e-07 Length: 480
Score: 256.00 Matches: 120
Percent Similarity: 38.3% Conservative: 69
Best Local Similarity: 24.3% Mismatches: 181
Query Match: 9.0% Indels: 124
DB: 2 Gaps: 18
US-10-665-990A-13 (1-1561) x AH2333 (1-480)
QY 89 TCTTCATGTTGTCGCCCTCAGAGACGGAGCAAGCGCTCATTTCAATACTTCCAAA 148
DB 98 SerAlaLysLeuGluProLeuGln----- 105
QY 149 CTGTGCTCTGGACAACATCTGCAAAATCGGCACACCCCTCATATAACAACGGGCTATCC 208
DB 106 ---LeuLeuAlaGluAlaPheIleGlyIleProPheThrSerGlyAsnAla----- 122
QY 209 GACATCTACCTGCTCGACGACCCCGCAAGCCCTTTCGCGCGCGCGCCCTTATCGAA 268
DB 123 -----LysLeuLeuIleAsnGlyGlnGlnThrTyrAlaAlaMetLeuSerAlaIleAla 140
QY 269 TCTGCGGAACACAGCTCGATTTCGAATCTACATTTGGCGCAACGACATTCGCGCAGG 328
DB 141 SerAlaAsnSerTyrIleLeuLeuGlnSerTyrIleValValAspLysAlaGlyAsn 160
QY 329 CTGCTGTTCAACCTCATCTGCTTTCGCGCAAGCGCGGTGCGGTACGCTGCTGTG 388
DB 161 GluPheLysAspAlaLeuIleAlaLysAlaLysGlnGlyIleArgValTyrLeuIleTyr 180
QY 389 GACGACAAACAACGCGCGGTGTGACGATCTCTGCTGCGCCCTCGACGAGCCATCCCAAT 448
DB 181 AspGluIleGlySerAsnLysIleSerArgLeuTyrValLysSerLeuGlnLysTyrAsp 200
QY 449 ATCGAAGTCGCGCTGTTCAACCCCTTCCTCTAGCAAAATGGCGCGCACTCGGCTACCTG 508
DB 201 IleGlnValSerAlaPheHisThrThrArgGlyArgGlyAsnArgPhe----- 216
QY 509 ACCGACTTCCCGCTCAACCGCGCATGCACAAATCCTTTACCGCGCAACCGC 568
```

Db 217 -----GlnLeuAsnPheArgAsnHisArgLysIleLeuValValAspGlyArg 232
QY GCCACCATACCTCGCGGACGCAATATCGCGCAGCAATCTTCAAAGTCGGTGAGAC--- 625
Db 233 ThrAlaPheIleGlyLeuAsnIleSerAspGlyArg-----LeuGlyLysAsnPro 250
QY 626 -----ACCGTTTTCGCGACTCGACATCTCCGCCCGCGCGCGTCTGCGGCGAAGTA 679
Db 251 ArgLeuSerProTyrArgAspThrHisMetMetLeuGluGlyProThrValGlnSerLeu 270
QY 680 TCGCAGCACTTC-----GACCGGTACTGGCAGAGCCATTCGCCGCCACAGCCACCGCGC 733
Db 271 GlnGlyCysPheLeuGlnAspTyrTrpAlaThr----- 282
QY 734 ATCATCCGCGCGGCAACATCGGCAAGGCTCTTCAAGCACTCGGATACACAGCAAGAAACA 793
Db 282 ----- 282
QY 794 TCCAGACACGGCTCTCGCTACCGGCAAAACCGTCGAACAGTCGCCCTCTACCAAAAA 853
Db 283 -----ArgGlnValIleAspValAsnTrpGlnValGlnPro 294
QY 854 ATACAGACGGACGCATCGACTCGCAGAGCGTCCAAACCCGCCCTGATCAGGACACC--- 910
Db 295 -----AsnTrpGluSerAspTyrThrAlaLeuValPheProThrGly 308
QY 911 CCTGCAAAAGACTCGACCGCGACCGCGCAACCGCGGATTCGCCGGAGGCTG----- 964
Db 309 ProAla-----AspLysLeuLys-----AlaCysLysLeuPhePhe 320
QY 965 CAAGACGGCTCAACAGCGCGGAAAGCGTCTATCTGGTTTCACCTATTTCTGTCCTCT 1024
Db 321 ValSerAlaIleAsnGlnAlaGlnThrArgLeuTrpIleAlaThrProTyrPheValPro 340
QY 1025 ACAAAATCCGGCAGCAGCTCGCAAACTGTGTGAGGACGCATAGACGTACCGTC 1084
Db 341 AspAspSerThrLeuThrAlaLeuLysLeuAlaLeuArgGlyValAspValArgIle 360
QY 1085 CTGACCAACTCGCTACAGGGACCGAGCTTGC CGCGCTTCATTCGGCTACGTCAAATAC 1144
Db 361 IleLeuProAsn-----ArgProAspHisLeuLeuValTyrLeuCysSerPheSerTyr 378
QY 1145 CGAAACCGGTGTCAAAGCGCGCAAACTCTACGAGCTGCAACCCCAACCATCGCGTC 1204
Db 379 TyrThrGluMetLysAlaThrAsnIleLysLeuTyrArgTyrLys-----HisGlyPhe 396
QY 1205 CCGCCACAAAAGCAAGCGCTGACCGGAGCTCCGTAACCGCTCGCTGATGCCAAACC 1264
Db 397 -----MetHisGlnLysVal 401
QY 1265 TTCATTGTGACGGCAACGCATCTTCATCGGCTCATTCAACTCGACCCCGTTCGCA 1324
Db 402 IleLeuIleAspLysGluMetAlaGlyValGlyThrValAsnLeuAspAsnArgSerPhe 421
QY 1325 CGGCTCAATACCGAA-----ATGGCGGTCTCATCGAAAGCCCAAAATFCGAGAACAGATG 1381
Db 422 PheLeuAsnPheGluValMetGlyPheValAlaAsnSerGlnPheValLysSerValGlu 441
QY 1382 GAGGGACCTCGCGATACACACCGAATAGCCTACCGCTTACCGCTTCGACAAACAC 1441
Db 442 GluMetLeuGlnAlaAspLeu-----LysAlaAlaLeuAlaValAspPheSerAspTyr 459
QY 1442 AACCGCTCAATGGCAGCATCCGCGCACCGGAAACCTTACCGCAAGCAACCGAAGCC 1501
Db 460 GluArg-----Lys 462
QY 1502 AAATCTTGGAAACGATCGCGGCAAAATCTATCCCTGTGTG 1543
Db 463 TyrLeuTrpPheLysLeuAlaValArgIleSerSerLeuLeu 476
RESULT 39
T43481
probable mucin DKFZp434C196.1 - human (fragment)

N;Alternate names: protein DKFZp434B0635.1
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43481; T34549; T17264
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z22514
A;Accession: T43481
A;Molecule type: mRNA
A;Residues: 1-580 <AAA>
A;Cross-references: UNIPROT:Q9UF83; UNIPARC:UPI000006D400; EMBL:AL133561; NID:G6599133;
A;Experimental source: adult testis; clone DKFZp434C196
R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, October 1999
A;Reference number: Z21540
A;Accession: T34549
A;Molecule type: mRNA
A;Residues: 262-580 <POU1>
A;Cross-references: UNIPARC:UPI0000070F36; EMBL:AL122069; NID:G6102864; PIDN:CAB59245.2
A;Experimental source: adult testis; clone DKFZp434B0635
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18723
A;Accession: T17264
A;Molecule type: mRNA
A;Residues: 262-580 <POU2>
A;Cross-references: UNIPARC:UPI0000070F36; EMBL:AL117481; NID:G5911958; PIDN:CAB55954.1
A;Experimental source: adult testis; clone DKFZp434B061
C;Genetics:
A;Note: DKFZp434C196.1; DKFZp434B0635.1; DKFZp434B061.1
Alignment Scores:
Pred. No.: 2,37e-07 Length: 580
Score: 252.50 Matches: 141
Percent Similarity: 37.4% Conservative: 53
Best Local Similarity: 27.2% Mismatches: 219
Query Match: 8.9% Indels: 108
DB: 2 Gaps: 22
US-10-665-990A-13 (1-1561) x T43481 (1-580)
QY 154 CCTCTTGACCAATCATCT-----GCAATCCGGCAGCACCCCTCATAA 195
Db 82 ProProArgAlaSerProThrArgLysProProArgAlaSerProArgThrProSerArg 101
QY 196 CAACGGCTATCGACATCTACCTGCTCGACGACCCCGACAGCCCTTG----- 245
Db 102 AlaSerProThrArgArgLeuProArgAlaSerProMetGlySerProHisArgAlaSer 121
QY 246 -----CGCCCGCGCGCTTATCAATCTGCCGACAGCCTCGATTTCG 293
Db 122 ProMetArgThrProProArgAlaSerProThrGlyThrProSerThrAlaSerProThr 141
QY 294 AATPACTATTTGGCGCAACGACATTTCCGGCAGGCTGC-----TGTTCACCTCA 344
Db 142 GlyThrProSerSerAlaSer-----ProThrGlyThrProProArgAlaSerProThr 159
QY 345 TGTACCTTCCGCGAGAACCGCGGTGCGCTGCTGCTGTTGGACGACCAACACACGC 404
Db 160 GlyThrProPro-ArgAlaTrpAlaThrArgSerProSerThr-----Al 174
QY 405 GCGGGTTGGACGATCTCTGCTCGC-----CCTCGACAGCCATCCCAATATCGAAG 455
Db 174 aSerLeuThrArgThrProSerArgAlaSerLeuThrArgTrpProProArgAlaSerPr 194
QY 456 TGGCGCTGTTCACCCCTTGTCTTACGAAATGGCGCGCAGCTCGGCTACCTGACCGACT 515
Db 194 oThrArgThrProProArgGluSerProArgMetSerHisArgAlaSerProThrArgTh 214
QY 516 TCCCCCGCTCAACCCCGCATGCACAAATCTTTTACCGCGGACCAACCGCGCCACCA 575
Db 214 rProProArgAlaSerPro-----ThrArgArgProProArgAlaSe 228

```
QY 576 TACTCGCGGACGCAATATCGCGGACGAAATCTTCAAAGTCGGTGAGGACACCGTTT-TC 634
Db 228 rProthrArgThrProProArgGluSerLeu::: ||| ||| ||| ||| ||| |||
QY 635 GCGGACCTGGACATCTCCGCCACCGGACGTCGTCGGGAGAGTATCGCACGACTTCGAC 694
Db 245 rPro-----ThrArgMetProProArgAlaSerProThrArgProProArgAlaSe 263
QY 695 CGCTACTGGG-----CAAGCATTCCGCGCCACAAACGCGACCGCATCATCCGACG 745
Db 263 rProThrGlySerProProArgAlaSerProMetThrProProArgAlaSerProArgTh 283
QY 746 GGCACATCGGCAAGGCTTCAACGACTCGGATACAGGACGAAACATCGACACGCG 805
Db 283 rProProArgAla-----SerProThrThrThr-----ProSerAr 295
QY 806 CTCTCGCGCTACCGGAAACCGTCGAACAGTCGCGCCCTCTACCAAAAAATACAGA---- 860
Db 295 gAlaSerLeuThrArgThrProSerTrpAlaSerProThrThrThrProSerArgAlaSe 315
QY 861 -----CGGAGCGCATCGAC 874
Db 315 rLeuMetLysMetGluSerThrValSerIleThrArgThrProProArgAlaSerProTh 335
QY 875 TGGCAGAGGTCCAAACCGGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCAC 934
Db 335 rGlyThrProSerArgAlaSerProThrGlyThrProSerArgAlaSerLeuThrGlySe 355
QY 935 GCGCCCAACCGCCGATTCCGCGGAGGCTGCAAGCGCGCTCAACACAGCCGCAAAAAAGC 994
Db 355 rProSerArgAlaSerLeuThrGly-----ThrProSerArgAl 368
QY 995 GTCTATCTGGTTTCAACCTATTTCGTCCTCAAAAATCGGCACAGACGCACTGGCAAAA 1054
Db 368 aserLeu-----IleGlyThrProSerArgAla----- 377
QY 1055 CTGGTGACGAGCGCATAGACGTTACCGCTCTGACCAACTCGCTACAGCGCAGCACGTT 1114
Db 378 -----SerLeuIleGlyThrProSerArgAlaSerLeuThrGlyTh 391
QY 1115 GCGCGCG-----TCCATTTCGCGTACGTCAAAT-----ACCGAAAA 1150
Db 391 rProProArgAlaSerLeuThrGlyThrSerSerThrAlaSerLeuThrArgThrProSe 411
QY 1151 CGCGTGTCAAAGCGGCGATCAAACTCTACGAGCTGCAACCAACCAATCGCTCCCGCC 1210
Db 411 rArgAlaSerLeuThrArgThrGlnSerSerSerSerLeuThrArgThrProSerMetAl 431
QY 1211 ACAAAAGACAAGGCTGACCGGACGCTCCGTAACCGCTGCATGCCAAACCTTCATT 1270
Db 431 aserLeuThrArgThrProProArgAlaSerLeuThrArgThrProProArgAlaSerLe 451
QY 1271 GTGGAGCGCAACGCAATCTTCATCGCTCATTCACACCTCGACCCCGCTTCGCGCGCTC 1330
Db 451 u-----ThrArgThrProProArgAlaSerLeuThrArgThrPro---ProArgAlaSe 468
QY 1331 AATA-----CCGAAATGGCGTGTGCATCGAAAGCCCAAAATCGCAGAACAGATG 1381
Db 468 rLeuThrArgThrProSerMetValSerLeuLysArgSerProSerArgAlaSerLeuTh 488
QY 1382 GAGCGCACCTCGCGCATACACACCCGAAATACGCTACC-----GGTTACCTCGAC 1435
Db 488 rArgThrProSerArgAlaSerLeuThrMetThrProSerArgAlaSerLeuThrArgTh 508
QY 1436 AAACACAAACCGCTCGCAATGGCAGCATCCCGCCACCCGCAAAAAACCTACCCGAAACGCC 1495
Db 508 rProSerThrAlaSerLeuThrGlyThrProProThrAlaSerLeuThrArgThrProPr 528
QY 1496 GAAGCCAAACTTTGGAACCGCATCGCGCAAAAAATCTATCCTCGCTGCCCATCG 1550
Db 528 othrAlaSerLeuThrArgSerProProThrAlaSerLeuThrArgThrProSer 546
```

RESULT 40

```
S21054
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain - human
N:Alternate names: DNA-directed RNA polymerase B largest chain
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S21054; S18987
R:Wintzerith, M.; Aker, J.; Vicaire, S.; Vigneron, M.; Kedinger, C.
Nucleic Acids Res. 20, 910, 1992
A:Title: Complete sequence of the human RNA polymerase II largest subunit.
A:Reference number: S21054; MUID:92178992; PMID:1542581
A:Accession: S21054
A:Molecule type: mRNA
A:Residues: 1-1970 <WIN>
A:Cross-references: UNIPROT:P24928; UNIPARC:UPI000000037D; EMBL:X63564; NID:g36123; PIDN
C:Genetics:
A:Gene: GDB:POLR2A; POLRA
A:Cross-references: GDB:120306; OMIM:180660
A:Map position: 17p13.1-17p13.1
C:Superfamily: human DNA-directed RNA polymerase II largest chain
F:Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc fing
F:71-87/Region: zinc finger CCHH motif

Alignment Scores:
Pred. No.: 3,4e-07 Length: 1970
Score: 250.50 Matches: 123
Percent Similarity: 38.2% Conservative: 69
Best Local Similarity: 24.5% Mismatches: 196
Query Match: 8.8% Indels: 115
DB: 1 Gaps: 24

US-10-665-990A-13 (1-1561) x S21054 (1-1970)
QY 108 TGGAAAGACGAGCGAAGCCGCTATTTCATATCTTCCAAACCTGCTCCTCTGGACAACA 167
Db 1520 TrpAsnGlnGly-----AlaThrProAlaTyxGlyAlaTrpSerPro 1533
QY 168 TCCTGCAAAATCGGCACACCCCTCATATAACAGGGCTATCCGACATCTACCTGC----- 221
Db 1534 SerValGlySerGlyMetThrProGlyAlaAlaGlyPheSerProSerAlaAlaSerAsp 1553
QY 222 -----TCGACGACCCCGACGAGCCCTTG 245
Db 1554 AlaSerGlyPheSerProGlyTyxSerProAlaTrpSerProThrProGlySerProGly 1573
QY 246 CGCGCCGCGCGCCCTTATCGAATCTGCGAACAACAGCCTCGATTGCAATACTACATT 305
Db 1574 SerProGlyProSerSerProTyxIleProSer----- 1584
QY 306 GCGCAACGACATTTTCGCGCAGGCTGTGTTCAACCTCATGTACCTTGC CGCAGAACGCG 365
Db 1585 -----ProGlyGlyAlaMetSerProSerTyxSer---ProThrSerPro 1598
QY 366 GCGTGC GGTACGCTGTGTTGGACGACAAACACGCGCGGGTTGGACGATCTCTGTC 425
Db 1599 Ala-----TyrGluProArgSerProGlyGlyTyxThrProGln--- 1611
QY 426 TCGCCCTCGACAGCATCCCAATATCGAAGTCGCGCTGTTCACCCCTTCGCTCTACGCA 485
Db 1612 SerProSerTyxSer---ProThrSerProSerTyxSerProThrSerProSerTyxSer 1630
QY 486 -----AATGCGCGCACTCGCTACCTGACCGACT---TCGCCCGCTCAACCGCGCA 536
Db 1631 ProThrSerProAsnTyxSerProThrSerProSerTyxSerProThrSerProSerTyx 1650
QY 537 TGCACAAATATCTTTTACC CGCGCAACACCGCGCCACCATACTCTCGCGGACCAATATCG 596
Db 1651 SerProThrSerProSerTyxSerProThrSerProSerTyxSerProThrSerProSer 1670
QY 597 GCGACCAATACTTCAAAGTCGGTAGGACACCGTTTTCGCCGACCTCGACATCTCTCGCA 656
Db 1671 TyrSerProThrSerProSerTyx---SerProThrSerProSerTyxSerProThrSer 1689
QY 657 CGGCGACGCTCG-----TCGCGCAAGTATCGACGACTTCGACCGCTACTGGCAAGCC 710
```



```
Db 1690 ProSerTyrSerProThrSerProSerTyrSerProThrSerProThrSerProThr 1709
Qy 711 ATTCCGCCACAAACCGCCGATCATCCGCGGCAACATCGGCAAGGCTTCAAG 770
Db 1710 SerProSerTyrSerProThrSerProSerTyrSerProThrSerProSer 1726
Qy 771 CACTCGGATACAGACGAGAAACATCCAGACAGCGCTCTCGCTACCGCGAAACCGTCG 830
Db 1727 TyrSerProThrSerProSerProThrSerProThrSerProThrSerPro 1746
Qy 831 AACAGTCGCCCTCTACCAAAATAACAGAGGAGCGATCGACTGGCAGAGCGTCCAA 890
Db 1747 AsnTyrSerProThrSerProAsnTyrThrProThrSerProSerTyrSerProThrSer 1766
Qy 891 CCCGCTCATGAGGACACCCCTGCAAAAGGACTCGACCGCGCGCAACCGCGCA 950
Db 1767 Pro-----SerTyrSerProThrSerProAsnTyrThrProThrSerProAsn----- 1782
Qy 951 TTGCGGGAGGCTGCAAGACGCGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTCAC 1010
Db 1783 -----TyrSerProThrSerProSer-----TyrSer 1791
Qy 1011 CCTATTTCGTCCTACAAAATCCGGCACAGACGCACTGGCAAAACTGGTGCAGGACGGCA 1070
Db 1792 ProThrSerProSerTyrSerProThrSerProSerTyrSer----- 1805
Qy 1071 TAGAGTTACCGTCTGACCAACTCGCTAC-----AGCGACGCGCTTCCCGCGCTCC 1124
Db 1806 -----ProSerPro-----ArgTyrThrProGlnSerProThrTyrThrProSer 1821
Qy 1125 ATTCGGCTAGGTCAAATACCGAAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGC 1184
Db 1822 SerProSerTyrSer-----ProSerSerProSerTyrSer 1833
Qy 1185 TGCACCCCAACGATCCGCTCCCGCCCAAAAGACAAAGGCGCTGACCGGACGTCCTGTA 1244
Db 1834 -----ProThrSerProLysTyrThr-----ProThrSerPro--- 1844
Qy 1245 CCAGCTGCATGCCAAACCTTCATTGTGGAGCGCAACGATCTTCATCGGCTCATTC 1304
Db 1845 -----SerTyrSer 1847
Qy 1305 ACCTCGACCCCGTTCCGACGCGTCAATACGAAATGGGCG---TCGTCAATCGAAAGCC 1361
Db 1848 ProSerSerProGluTyrThrProThrSerProLysTyrSerProThrSerProLysTyr 1867
Qy 1362 CCAAAATCGAGACAGATGGAGCGCACCTCGCGGATACACACCGGAATACCGCTACC 1421
Db 1868 SerProThrSerProLysTyrSerProThrSerProThrTyrSerProThrProLys 1887
Qy 1422 GCGTTACCTCGACAAACACACCGCTGCAATGGCAGATCCCGCCACCCGAAAAACCT 1481
Db 1888 TyrSerProThrSerProThrTyrSerProThrSerProValTyrThrProThrSerPro 1907
Qy 1482 ACCGGAACGAAACCGCAACCAAACTTTGGAAACGATCGCGCAAAATCCTATCCCTGC 1541
Db 1908 LysTyrSerProThrSerProThrTyrSerProThrSerProLysTyrSer---ProThr 1926
Qy 1542 TGCCCATCG 1550
Db 1927 SerProThr 1929
```

Search completed: May 2, 2006, 05:26:35
Job time : 143 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:09:35 ; Search time 41 Seconds
(without alignments)
1232.044 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAATLSLLPIEGLL 525

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2643	97.2	525	B81859	phospholipase D-fam
2	2573	94.6	508	B81083	cardiolipin syntha
3	1102.5	40.5	493	C64847	ymdC protein - Esc
4	1100.5	40.5	493	H90806	probable synthase
5	1093.5	40.2	493	D85666	conserved hypothet
6	1089.5	40.1	495	AE0636	hypothetical 55.9K
7	807	29.7	518	G97556	phospholipase D fa
8	807	29.7	518	AB2777	conserved hypothet
9	792	29.1	529	F82983	probable cardiolip
10	632	23.2	502	E71963	conserved hypothet
11	627	23.1	502	F64543	cardiolipin syntha
12	439	16.1	652	B82724	cardiolipin syntha
13	314	11.5	482	G70063	cardiolipin syntha
14	308	11.3	494	F90001	hypothetical prote
15	300	11.0	504	A11433	cardiolipin syntha
16	294.5	10.8	504	A11075	cardiolipin syntha
17	292	10.7	482	A11762	cardiolipin syntha
18	287	10.6	482	AG1387	cardiolipin syntha
19	284.5	10.5	505	G89906	hypothetical prote
20	282.5	10.4	481	B86744	cardiolipin syntha
21	276	10.1	503	B84007	cardiolipin synthe
22	275	10.1	398	C84125	cardiolipin synthe
23	273	10.0	476	T43863	cardiolipin syntha
24	270.5	9.9	467	H82711	cardiolipin syntha
25	268	9.9	490	B82971	cardiolipin syntha
26	266.5	9.8	500	S60089	cardiolipin syntha
27	263.5	9.7	492	AE3539	cardiolipin synthe
28	261	9.6	359	D83103	probable phospholi
29	260	9.6	510	E97307	probable cardiolip

ALIGNMENTS

RESULT 1

B81859

phospholipase D-family protein NMA1646 [imported] - Neisseria meningitidis (strain Z2491

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C;Accession: B81859

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; More

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandrea

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: B81859

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-525 <PAR>

A;Cross-references: UNIPROT:Q9JUT1; UNIPARC:UPI000005033F; GB:AL162756; GB:AL157959; N1

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA1646

C;Superfamily: cardiolipin synthase

Query Match	97.2%;	Score	2643;	DB 2;	Length	525;			
Best Local Similarity	97.3%;	Pred. No.	5.2e-182;						
Matches	511;	Conservative	5;	Mismatches	9;	Indels	0;	Gaps	0;
Qy	1	MHTDPKIQAMPSETISPMKTRSLISLLCLLLCCSSWLPPLEERTESRHNTPKPVLLDN	60						
Db	1	MRANPKTQAMPSETISLMKTRSLISLLCLLLCCSSWLPPLEERTESRHNTPKPVRLDN	60						
Qy	61	ILQIRHTPHNNGSLDIYLLDDPHEAALAAALIESAHSLSLQYIWRNDISGRLLFNLM	120						
Db	61	ILQIRHTPHNNGSLDIYLLNDPHEAFAARAALIESAHSLSLQYIWRNDISGRLLFNLV	120						
Qy	121	YLAERGVRRVRLDDNNTRGLDLDLLDSDHNIEVRLNFPVLRKWRALGYLTDPRRL	180						
Db	121	YLAERGVRRVRLDDNNTRGLDLDLLDSDHNIEVRLNFPVLRKWRALGYLTDPRRL	180						
Qy	181	NRNMENKSFADNRATILGGRNIGDEYFKVGVSDTVFADLDILATGSVVGSEVSHDFRYWA	240						
Db	181	NRNMENKSFADNRATILGGRNIGDEYFKVGVSDTVFADLDILATGSVVGSEVSHDFRYWA	240						
Qy	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVESQSPLYQKIQTGRIDWQSV	300						
Db	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVESQSPLYQKIQTGRIDWQSV	300						
Qy	301	QTRLSIDTTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDLAKLVQD	360						
Db	301	QTRLSIDTTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDLAKLVQD	360						
Qy	361	GIDVTVLTNSLQATDVAHVHSGYKVRKPLLKAGIKLYELQPNHVPATPKDKGLTGSSVT	420						
Db	361	GIDVTVLTNSLQATDVAHVHSGYKVRKPLLKAGIKLYELQPNHVPATPKDKGLTGSSVT	420						

cardiolipin syntha
cardiolipin syntha
cardiolipin syntha
cardiolipin synthe
probable phospholi
cardiolipin syntha
cardiolipin syntha
cardiolipin syntha
probable phospholi
ybhO protein - Esc
probable synthetas
cardiolipin synthe
cardiolipin synthe
cardiolipin syntha
cardiolipin syntha
cardiolipin synthe
cardiolipin synthe

QY 421 SLHAKTFIVDGKRIFIGSFNLDPRGARLNTENGVVIESPKIAEQMERTLADTTPEYAYRV 480
|||||
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRGARLNTENGVVIESPKIAEQMERTLADTSPYAYRV 480
|||||
QY 481 TLDKHNRLQWHDPAATKTYNEPEAKLWKRIAATKILSLPIEGLL 525
|||||
Db 481 TLDNRHLQWHDPAATKTYNEPEAKLWKRIAATKILSLPIESLL 525
|||||
RESULT 2
B81083
cardiolipin synthetase family protein NMB1434 [imported] - Neisseria meningitidis (strain
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81083
R:Tetelin, H.; Saunders, N.J.; Heidelsberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81083
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-508 <TET>
A:Cross-references: UNIPROT:Q9JYU0; UNIPARC:UPI00000C46B9; GB:AE002494; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1434
C:Superfamily: cardiolipin synthase
Query Match 94.6%; Score 2573; DB 2; Length 508;
Best Local Similarity 97.8%; Pred. No. 5.2e-177;
Matches 497; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 18 MKTRSLISLLCLLSCSSWLPPLBERTESRHFNTPSKVRLDNLILQIRHTPHNGLSDIY 77
Db 1 MKTRSLISLLCLLSCSSWLPPLBERTESRHFNTPSKVRLDNLILQIRHTPHNGLSDIY 60
|||||
QY 78 LDDPHEALARAALIESAHSLDLQYIWRNDISGRLLFNLMYLAARGVVRLLDDN 137
Db 61 LNDPHEAFARAALIESAHSLDLQYIWRNDISGRLLFNLMYLAARGVVRLLDDN 120
|||||
QY 138 NTRGLDDLLALDHPNIEVRLFNPFVLRKWRALGYLTFDPRLNRRMNKGFADNRATI 197
Db 121 NTRGLDDLLALDHPNIEVRLFNPFVLRKWRALGYLTFDPRLNRRMNKGFADNRATI 180
|||||
QY 198 LGGRNIGDEYFKVGEDTFADLIDLATGVSVEVSHDFDRYASHAHNATRIIRSGNIG 257
Db 181 LGGRNIGDEYFKVGEDTFADLIDLATGVSVEVSHDFDRYASHAHNATRIIRSGDIG 240
|||||
QY 258 KGLQALGYNDETSRRALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLSIDTPAKGLDRDR 317
Db 241 KGLQALGYNDETSRRALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLSIDTPAKGLDRDR 300
|||||
QY 318 RKPIAGRLQDALQKPEKSVLYSPFYVPTKSGTGDALAKLVQDGDIDVTVLTNSLQATDVA 377
Db 301 RKPIAGRLQDALQKPEKSVLYSPFYVPTKSGTGDALAKLVQDGDIDVTVLTNSLQATDVA 360
|||||
QY 378 AVHSGYVKYRPELLKAGIKLYELOQNHAVPATKDKGLTGSSVTSIHKATFIVDGKRIEIG 437
Db 361 AVHSGYVKYRPELLKAGIKLYELOQNHAVPATKDKGLTGSSVTSIHKATFIVDGKRIEIG 420
|||||
QY 438 SFNLDPGARLNTENGVVIESPKIAEQMERTLADTTPEYAYRVTLDDKHNRLQWHDPAATRK 497
Db 421 SFNLDPGARLNTENGVVIESPKIAEQMERTLADTTPEYAYRVTLDDKHNRLQWHDPAATRK 480
|||||
QY 498 TYPNEPEAKLWKRIAATKILSLPIEGLL 525
Db 481 TYPNEPEAKLWKRIAATKILSLPIEGLL 508
|||||

RESULT 3
C64847
ymdC protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: C64847
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64847
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <BLAT>
A:Cross-references: UNIPROT:P75919; UNIPARC:UPI000013B993; GB:AE000206; GB:U00096; NID:3
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ymdC
C:Superfamily: cardiolipin synthase
Query Match 40.5%; Score 1102.5; DB 2; Length 493;
Best Local Similarity 47.6%; Pred. No. 2.2e-71;
Matches 221; Conservative 83; Mismatches 143; Indels 17; Gaps 5;
QY 72 GLSDIYLLDDPHEALARAALIESAHSLDLQYIWRNDISGRLLFNLMYLAARGVVR 131
Db 37 QCCGLFPLEKSLDAFAARYLAENAEHTLDVQYVIWQDDMSGRLLFSAALLAAKRGVVR 96
|||||
QY 132 LLLDDNTRGLDLLLLALDHPNIEVRLFNPFVLRKWRALGYLTFDPRLNRRMNKSFPA 191
Db 97 LLLDDNTRGLDLLLLALDHPNIEVRLFNPFVLRKWRALGYLTFDPRLNRRMNKSFV 156
|||||
QY 192 DNRTILGGRNIGDEYFKVGEDTFADLIDLATGVSVEVSHDFDRYASHAHNATRII 251
Db 157 DGVTVLGGRNIGDAYFCAGEEPLFSDLDWAIGPVVEDVADDFARYWYCKSVPLQQLV 216
|||||
QY 252 RSGNIGK-----LQALGYNDETSRRALLRYRETVEQSPLYQKIQTGRIDWQSVQTRL 304
Db 217 ---DVPEGMADRIELPASHNDAMTH----RYLRKMESSPFINHLVDGTPLIWAKTRL 269
|||||
QY 305 ISDTPAKGLDRDRKPPKPIAGRLQDALQKPEKSVLYSPFYVPTKSGTGDALAKLVQDGDIV 364
Db 270 LSDDPKAGEGKAKRHSLLPQRLFDIMGSPSERIDIISYFVPTTRAGVAQLLRWVRKGVKI 329
|||||
QY 365 TVLTNSLQATDVAAVHSGYVKYRPELLKAGIKLYELOQNHAVPAT-KDKGLTGSSVTSLSH 423
Db 330 ALLTNSLAANDVAVVHAGYARWRKLLRYGVVELYELKPTREQSSTLHDRGITGNSGSLH 389
|||||
QY 424 AKTIVDGKRIFGSFNLDPRGARLNTENGVVIESPKIAEQMERTLADTTPEYAYRVTLDD 483
Db 390 AKTFSIDGKTVFIGSFNLDPRGARLNTENGVVIESPKIAEQMERTLADTTPEYAYRVTLDD 449
|||||
QY 484 KHNRLQWHDPAATK-TYPNEPEAKLWKRIAATKILSLPIEGLL 525
Db 450 RWGRINWDRHAKKEIILKKEPATSFVKRVVVRVLAASILPVEWLL 493
|||||
RESULT 4
H90806
probable synthase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90806
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90806
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-493 <HAY>
A;Cross-references: UNIPROT:Q8X9I7; UNIPARC:UPI00000D07FD; GB:BA000007; PIDN:BA034847.1;
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: ECs1424
C;Superfamily: cardiolipin synthase

Query Match 40.5%; Score 1100.5; DB 2; Length 493;
Best Local Similarity 47.4%; Pred. No. 3e-71;
Matches 220; Conservative 84; Mismatches 143; Indels 17; Gaps 5;

```
QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDLYIWRNDISGRLLFNLMYLAERGVRVR 131
Db 37 GQGLFPLEKSLDFAFARYLAEMSEHTLDVQYIWDMSGRLLFSALLAAAKRGVRVR 96
QY 132 LLLDDNNTRGLDDLLALDHPNIEVRLFNFPVLRKWRALGYLTDFFRLNRRMHNKSF 191
Db 97 LLLDDNNTPGLDDILRLDSDHPRIEVLRFNFPFSLRLLRPGLYITDFSLNRRMHNKSF 156
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDPDRYWAHSHAHNATRII 251
Db 157 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVEDVADDFARYWYCKSVSPLOQVL 216
QY 252 RSGNIGK-----LQALGVNDTSRHALLRYRETVEQSPLYQKIQTGRIDWOSVQTRL 304
Db 217 ---DVPEGEMADRIELPASWHNDAMTH----RYLRKMESPFFINHLVDGTLPLIWAQTRL 269
QY 305 ISDTPAKGLDRDRKPPKIAGRLQDALQKPEKSVVLSVFPVPTKSGTDLAKLVQDGDIV 364
Db 270 LSDDPKAGEKAKHSLLPQRLFDIMGSPSERDIISSYFVPTAGVAQLLRVMVRKGVKI 329
QY 365 TVLTNSLQATDVAHVSHGVYKRYKPLKAGIKLYELOPNHAVPAT-KDKGLTGSVSTSLH 423
Db 330 AILTNSLAANDVAVVHAGYARWRKLLRYGVELYELKPTREQSSTLHDTGNSGASLH 389
QY 424 AKTFIVDGKRIFTGSFNLDRPSARLNTMGVIVIESPKIAEQMERTLADTTPEYAYRVTL 483
Db 390 AKTFSIDGKTVFIGSFNDFPRSTLLNTMGFVIESETLAQLIDKRFQSQYDAAWQLRLD 449
QY 484 KHNRLQWHDPATRK--TYPNEPEAKLWKRIAATKLSLLPIEGLL 525
Db 450 RWGRINWVDRHAKKEIVLKEPATSFWKRVWVRILASILPVEWLL 493
```

RESULT 5
D85666
probable synthase ymdC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85666
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85666
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: UNIPROT:Q8X9I7; UNIPARC:UPI0000165766; GB:AE005174; NID:g12514578; H
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ymdC
C;Superfamily: cardiolipin synthase

Query Match 40.2%; Score 1093.5; DB 2; Length 493;
Best Local Similarity 47.2%; Pred. No. 9.6e-71;
Matches 219; Conservative 84; Mismatches 144; Indels 17; Gaps 5;

```
QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDLYIWRNDISGRLLFNLMYLAERGVRVR 131
Db 37 GQGLFPLEKSLDFAFARYLAEMSEHTLDVQYIWDMSGRLLFSALLAAAKRGVRVR 96
```

```
QY 132 LLLDDNNTRGLDDLLALDHPNIEVRLFNFPVLRKWRALGYLTDFFRLNRRMHNKSF 191
Db 97 LLLDDNNTPGLDDILRLDSDHPRIEVLRFNFPFSLRLLRPGLYITDFSLNRRMHNKSF 156
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDPDRYWAHSHAHNATRII 251
Db 157 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVEDVADDFARYWYCKSVSPLOQVL 216
QY 252 RSGNIGK-----LQALGVNDTSRHALLRYRETVEQSPLYQKIQTGRIDWOSVQTRL 304
Db 217 ---DVPEGEMADRIELPASWHNDAMTH----RYLRKMESPFFINHLVDGTLPLIWAQTRL 269
QY 305 ISDTPAKGLDRDRKPPKIAGRLQDALQKPEKSVVLSVFPVPTKSGTDLAKLVQDGDIV 364
Db 270 LSDDPKAGEKAKHSLLPQRLFDIMGSPSERDIISSYFVPTAGVAQLLRVMVRKGVKI 329
QY 365 TVLTNSLQATDVAHVSHGVYKRYKPLKAGIKLYELOPNHAVPAT-KDKGLTGSVSTSLH 423
Db 330 AILTNSLAANDVAVVHAGYARWRKLLRYGVELYELKPTREQSSTLHDTGNSGASLH 389
QY 424 AKTFIVDGKRIFTGSFNLDRPSARLNTMGVIVIESPKIAEQMERTLADTTPEYAYRVTL 483
Db 390 AKTFSIDGKTVFIGSFNDFPRSTLLNTMGFVIESETLAQLIDKRFQSQYDAAWQLRLD 449
QY 484 KHNRLQWHDPATRK--TYPNEPEAKLWKRIAATKLSLLPIEGLL 525
Db 450 RWGRINWVDRHAKKEIVLKEPATSFWKRVWVRILASILPVEWLL 493
```

RESULT 6

AE0636

conserved hypothetical protein STY1185 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004
C;Accession: AE0636
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A;title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE0636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-495 <PAR>
A;Cross-references: UNIPARC:UPI0000059FC8; GB:AL513382; PIDN:CAD08272.1; PID:g16502319;
C;Genetics:
A;Gene: STY1185
C;Superfamily: cardiolipin synthase

Query Match 40.1%; Score 1089.5; DB 2; Length 495;
Best Local Similarity 47.4%; Pred. No. 1.9e-70;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

```
QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDLYIWRNDISGRLLFNLMYLAERGVRVR 131
Db 39 GEGCILALDNLSDAFARYLRTMAAATLDVQYIWEEDMSGRLLFSALLAAAKRGVHR 98
QY 132 LLLDDNNTRGLDDLLALDHPNIEVRLFNFPVLRKWRALGYLTDFFRLNRRMHNKSF 191
Db 99 LLLDDNNTPGLDDTLRLDSDHPNIEVRLFNFPFSLRLLRPGLYITDFARLNRHVNKSITA 158
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDPDRYWAHSHAHNATRII 251
Db 159 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVDVANDFERWYRWCSSVSTLQOVL 218
QY 252 RSGNIGKGLQAL-----GYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWOSVQT 302
Db 219 SLSE-----QELTORIELPESWYNDEITR----RYLHKLETQSFMAIDLDRGTPLIWAQT 269
QY 303 RLISDTPAKGLDRDRKPPKIAGRLQDALQKPEKSVVLSVFPVPTKSGTDLAKLVQDGI 362
```

Db 270 RLLSDSKGEGKAQRHSLLPQRLFDVWGSGSTERIDIISAYFVTRAGVAQLNLVRKGV 329
Qy 363 DVTVLTNSLQATDVAAVHSGVKKYKPKLLKAGIKLYELQPNHA-VPA TKOKGLGTGSSVTS 421
Db 330 KIAITNSLAANDVAHVHAGYARWRKKLLRYGVVELYELKPTREHETAVHDRGLTNGSGSS 389
Qy 422 LHAKTFIVDGKRIIFIGSFNLDPRARLNTENGCVVIESPKIAEQMERTLATDTPPYAYRVT 481
Db 390 LHAKTFIDSGSKVFGSLNFDPRSTLLNTENGFGVIESETLATLTHKRFQTSQRDAAWOLR 449
Qy 482 LDKHNRLQWHD--PATRKYTNEPEAKLWKRIAAKIILSLPIEGLL 525
Db 450 LDRNGRINWIDRQOEKEEKVLPKPEATRFQWRQVRLVLAAILPVEWLL 495
RESULT 7
G97556
hypothetical 55_9K protein in csgc-mdog intergenic region [imported] - Agrobacterium tum
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97556
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97556
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI000000D1C0B; GB:AE007869; PIDN:AAK87408.1;
C:Genetics:
A:Map position: circular chromosome
C:Superfamily: cardiolipin synthase
Query Match 29.7%; Score 807; DB 2; Length 518;
Best Local Similarity 38.6%; Pred. No. 4e-50;
Matches 204; Conservative 85; Mismatches 196; Indels 44; Gaps 16;
Qy 23 LISLCLLLCSCSWLPPLERTESRHFNTSKPVLLDNIQLIRTPH--NNGLSD---IYL 78
Db 8 IILLMLIGPSLFVVGKQREKAIPKRPSTALPTEDETALDRHWQSRNGWNEKNALCL 67
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVVRVRLLLDDNN 138
Db 68 LHSNLDAPAVRAAARAAGRSGLDMYYMWNADLTGRLMMREVIAAADRGVVRVRLLLDDLG 127
Qy 139 TRGLDDLLALLDHPNTEVRLFNPFVLRK---WRALGYLTDFFPRLNRRMHNKSFADNRA 195
Db 128 VMSMDRIFPHADSHPNIELRFPNTRARENILHRSLELVLRFRSVNRMMHNKAWIADGRA 187
Qy 196 TILGRNIGDEYFKVGEDTVFADLDILATGSGVGEVSHDFRYWASHAHN-----AT 248
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDLAKLQVQGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYIFGKQGEVFEFLDLAERGV 355
Qy 249 RIIRSGNIGKGLQALGYNDTSRHALRYRETVEQSPLYQK---IQGRIDWQSVQTR-L 304
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDLAKLQVQGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYIFGKQGEVFEFLDLAERGV 355
Qy 363 DVTVLTNSLQATDVAAVHSGVKKYKPKLLKAGIKLYELQPNHNAVPA TKDKG---LTGSSV 419
Db 356 SLAILTNSLAATDVAHVHAGYARWRKKLLSGVRLHRL-----SQADQGSFTLRGSGQ 409
Qy 420 TSLHAKTFIVDGKRIIFIGSFNLDPRARLNTENGCVVIESPKIAEQMERTLATDTPPE-YAY 478
Db 410 ASLHTKAFTRDGETGYIGSLNFDPRSRASLNTMGVFNPSAPLVARMDRIFAEIERRTMSF 469

Qy 479 RVTLDKHNRLQW--HDPATRKTYNPEPEAKLWKRIAAKIILSLPIEGLL 525
Db 470 ELDIDSANRIVMWTEERGQPKIYRREPDAAISRIIAGIMRVLPLESQ 518
RESULT 8
AB2777
Phospholipase D family protein [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2777
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI000000D1C0B; GB:AE008688; PIDN:AAL42632.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atul630
A:Map position: circular chromosome
C:Superfamily: cardiolipin synthase
Query Match 29.7%; Score 807; DB 2; Length 518;
Best Local Similarity 38.6%; Pred. No. 4e-50;
Matches 204; Conservative 85; Mismatches 196; Indels 44; Gaps 16;
Qy 23 LISLCLLLCSCSWLPPLERTESRHFNTSKPVLLDNIQLIRTPH--NNGLSD---IYL 78
Db 8 IILLMLIGPSLFVVGKQREKAIPKRPSTALPTEDETALDRHWQSRNGWNEKNALCL 67
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVVRVRLLLDDNN 138
Db 68 LHSNLDAPAVRAAARAAGRSGLDMYYMWNADLTGRLMMREVIAAADRGVVRVRLLLDDLG 127
Qy 139 TRGLDDLLALLDHPNTEVRLFNPFVLRK---WRALGYLTDFFPRLNRRMHNKSFADNRA 195
Db 128 VMSMDRIFPHADSHPNIELRFPNTRARENILHRSLELVLRFRSVNRMMHNKAWIADGRA 187
Qy 196 TILGRNIGDEYFKVGEDTVFADLDILATGSGVGEVSHDFRYWASHAHN-----AT 248
Db 188 VIVGGRNIGDAYFDAERANFHDIFLGFGIVADATEIFDDYW--NSAVSPVRSLLAR 245
Qy 249 RIIRSGNIGKGLQALGYNDTSRHALRYRETVEQSPLYQK---IQGRIDWQSVQTR-L 304
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDLAKLQVQGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYIFGKQGEVFEFLDLAERGV 355
Qy 363 DVTVLTNSLQATDVAAVHSGVKKYKPKLLKAGIKLYELQPNHNAVPA TKDKG---LTGSSV 419
Db 356 SLAILTNSLAATDVAHVHAGYARWRKKLLSGVRLHRL-----SQADQGSFTLRGSGQ 409
Qy 420 TSLHAKTFIVDGKRIIFIGSFNLDPRARLNTENGCVVIESPKIAEQMERTLATDTPPE-YAY 478
Db 410 ASLHTKAFTRDGETGYIGSLNFDPRSRASLNTMGVFNPSAPLVARMDRIFAEIERRTMSF 469
Qy 479 RVTLDKHNRLQW--HDPATRKTYNPEPEAKLWKRIAAKIILSLPIEGLL 525
Db 470 ELDIDSANRIVMWTEERGQPKIYRREPDAAISRIIAGIMRVLPLESQ 518
RESULT 9
F82983

conserved hypothetical protein PA5310 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F82983
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Muzoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
gen
A:Reference number: A82950; MUID:20437337; PMID:10384043
A:Accession: F82983
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-529 <STO>
A:Cross-references: UNIPROT:Q9HTP4; UNIPARC:UPI00000C5F9D; GB:AE004943; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5310
C:Superfamily: cardiolipin synthase
Query Match 29.1%; Score 792; DB 2; Length 529;
Best Local Similarity 36.7%; Pred. No. 4.9e-49;
Matches 200; Conservative 86; Mismatches 189; Indels 70; Gaps 16;
QY 21 RSISLLCLLLCCSSWLPP-----LEERTESHFNSTKPVLLDNLILQIRHTP 68
Db 15 RRELLALLALSGCASTPPPPQSSALPAEGTWLARQAE-----IQGRDHP 59
QY 69 HNNGLSDIYLLDPHEALARAALIESAHSIDLOYYIWRNDISGRLLFNMYLAERG 128
Db 60 ---GQSGFHLLSASEDAFVARAALIRAAQSLDIQYIYVHDGLTTALAYELLKAADRGV 116
QY 129 RVRLLLDDNTRGLDLLLALDGHNPTEVLEFPVLRK---WRALGYLTDPPRLNRM 184
Db 117 RVRLDDTASDGDWYEIGVLSHPNLOVRLFNPLHGRATGTRGVGRFNLSQQHRRM 176
QY 185 HNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVEGSHDFRYWASHA 244
Db 177 HNKLLWLDGTAALVGGNLDGEYFNAPENMFTDLDLLGGPIANQLSHSFDQYWSAIS 236
QY 245 HNAT---RIIRSGNTGKGALQALGYNDTSRHALRY---RETVEQS-----PLYQKIQTGR 294
Db 237 RPIEDFLWRAPYGEEL-----ESARRKLQRYLRKESKESGYTRHFLDRGDQPR 285
QY 295 I-DWQS---VQTRLISDTPAKGLDRDRKPP--IAGRLQDALQKEKSVLYVSPYFVPT 347
Db 286 LGNWLNLNLTWARAEATWADPKVLSRGEPPHLLSPHLAAGLFGVQKEILLVSATFVPA 345
QY 348 KSGTDALAKLVQDGDIVTLTNSLOATDVAAHSGYVKYRKPILLKAGIKLYELQPNHVP 407
Db 346 KQGLNLTGKADSGVRVRLTNSLEATDVPVHAGVAPYRMALLEGVKLYELRANPDQP 405
QY 408 AT-KDKGLTGSSVTSLSHAKTFIVDGKRIFGSNLDPDRSARLNTENGWVIESPKIAEQME 466
Db 406 LSGAPWRLHGSSASSLHSAWVDRKRVFTGSGNFDPRSLWNTEVGVIVDPSPLAEQVR 465
QY 467 R-TLADTTPYAYRVTLDKHN---RLQWHD--PATRKYPEPEAKLWKRIAKILSLP 520
Db 466 QLALEGMAVSQVQVRIIDRSGSPKLVWIDRGRQAQVLNHEP-GSLWRRLNNAWAGMIG 524
QY 521 IEGLL 525
Db 525 LEKML 529
RESULT 10
E71963
probable cardiolipin synthase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: E71963
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat
hogen
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: E71963
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <ARN>
A:Cross-references: UNIPROT:Q9ZMP2; UNIPARC:UPI000013942E; GB:AE001456; GB:AE001439; NI
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0176
C:Superfamily: cardiolipin synthase
Query Match 23.2%; Score 632; DB 2; Length 502;
Best Local Similarity 32.8%; Pred. No. 1.4e-37;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;
QY 64 IRHTPHNGLSDIY-----LDDPHEALARAALIESAHSIDLDLQYYIWRN 109
Db 30 ISYDPTTTIGSLYAKNLKENPKHSAAILLEDGDFDALLHRVGLIRMSQKSIDMQTYIYKN 89
QY 110 DISGRLLFNMYLAERGVVRVRLLLDDNNTRGLD----DLLLALDHPNTEVRLFPVL 165
Db 90 DLSQVIAKELLNANRGVKVRIILDDN---GLDSDFSDIML-LNFHKNIEVKIFNFIYI 145
QY 166 RKWRALGY---LTDFFPLNRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDIL 222
Db 146 RN-KGURYFEMLADYERIKRMHKNLFIVDNFAVIIGGRNIGDYFDNDLDTNPLDLDAL 204
QY 223 ATGSVVGEVSHDFDRYVASHNAHNAIRIIS-----GNIGKGLQALGYNDTSRHAL 274
Db 205 FFGVASKAKESPENYWRFRSIPVS-LURTHKRLKNNVKEIAKLHEKIPISAEADANE 263
QY 275 LRYRETVESQPLYQ-KIQTRIDMQSVQTRLSIDTPAKGLDRDRKPPPIAGRIQDALQK 333
Db 264 KKVNDPIERFQKYQYPIYVG-----NAIFLADLPK-IDTPLYS-PIKIAFEKALKNA 314
QY 334 EKSVLYSVYFVPTKSGTDALAKLVQDGDIVTLTNSLOATDVAAHSGYVKYRKPILLKA 393
Db 315 KDSVFIASSYFIPGKKIMKIFKNQISKGIELNLTNSLSTDAIVVYGAWERYNRKLV 374
QY 394 GKLYELQPNHVAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFGSNLDPDRSARLNT 453
Db 375 GANVYIIRDFFNRQIKGR---FSTGKLSHGKTIIVFDLTLGSGNIDPRSAINTESA 431
QY 454 VIESPKIAEQMERTLADTTPYAYRVTLDKNRLQWHDPAKTKTY----PNPEAKLWK 509
Db 432 VLFDNPSFAKRVRLSLKDHQ-QQSWHLVLYRH-RVIWE--ATEGILIHEKNSPDTSF 487
QY 510 RTAAKILSLP 520
Db 488 RLKIKESKVL 498
RESULT 11
F64543
conserved hypothetical secreted protein HP0190 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: F64543
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I.
Nature 386, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64543
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-502 <TOM>
A:Cross-references: UNIPROT:P56117; UNIPARC:UPI000013942F; GB:AE000539; GB:AE000511; NI;

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: G70063

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-482 <KUN>

A:Cross-references: UNIPROT:P71040; UNIPARC:UPI0000060B8F; GB:Z99122; GB:AL009126; NID:9

A:Experimental source: strain 168

C:Genetics:

A:Gene: ywhE

C:Superfamily: cardiolipin synthase

Query Match 11.5%; Score 314; DB 2; Length 482;
Best Local Similarity 25.3%; Pred. No. 9.3e-15;
Matches 126; Conservative 65; Mismatches 189; Indels 118; Gaps 15;

Qy 43 ERTSRHFNTPSKVLLDNLQIRHTPHNGL-----SDIYLLDDPHEALAAALIESAE 97

Db 83 EDLETQKQFNRRATFNKDIYMLINNHAVFTEDNSVDITDGRKFORLLSDISKAK 142

Qy 98 HSLDLOYIWRNDISGRLLFNMYLAERGVRRLDDNNTRGLDALLALDHPNIEV 157

Db 143 DHIHLQYIYKGBELGKKLRDALLIQAKAGVQVRVLYDELGSRTLKKFKELREAGHV 202

Qy 158 RLEFPFVLRKWRALGYLTDPRINRPHNKSFTADNRATILGGRNIGDEYF-----KVGED 213

Db 203 EVFPFSKLRPINL-----RLNRYNRHKLVIIDGMTGYVGGFNVGDEYLGILNPKFG-- 252

Qy 214 TVFADLLDILATGTVVGVSHDFRYWASHSAHNATRIIRS-----GNIGKGLQALGY 265

Db 253 -YWRDTHIRLQGTAVHAIQTRFILDWQAASHHTLTYPHPDYGKPGNVGMQIVTSGP 311

Qy 266 NDETSRHALLRYRETVEQSPLYQKIOTGRIDWQSVQTRLSIDTPAKGLDRDRRKPIAGR 325

Db 312 DSE-----WEQIKNGYI-----KMISNA----- 329

Qy 326 LQDALKQPEKSVILVSPYFPTKSGTDALAKLVODGIDVTVLNLSQATDVAHVSHGYK 385

Db 330 -----KRSILIQTFYFPDASLLDALRIACLSGIDVNMIPN--KPDHAFVYWTLS 379

Qy 386 YRKPLKAGIKLYELOPNHAPVATKDKGLTSSVTSLSHAKTFIVDGKRIFGSFNLDPRS 445

Db 380 YIGDLLKAGATVY-----IYDNGF-----IHAKTIIVDDBIASVGTANIDVRS 422

Qy 446 ARLNTENGVVIESPKIAEQMERTLADTPPEYAYRVTLDKHNLQWHDPAIRK-TYPNEPE 504

Db 423 FRNFEVNAFIYDITIAKLVSIF-----KEDLL-----VSRKFTYEYELQ 463

Qy 505 AKLWKRIAKILSL-PI 521

Db 464 RPLWIRIKESVSRLLSPI 481

RESULT 14

F90001

hypothetical protein SA1891 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: F90001

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud

ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: F90001

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-494 <KUR>

A:Cross-references: UNIPROT:Q99SGS; UNIPARC:UPI00000D778F; GB:BA000018; PID:g13701883; H

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA1891

C:Superfamily: cardiolipin synthase

Query Match 11.3%; Score 308; DB 2; Length 494;
Best Local Similarity 23.2%; Pred. No. 2.6e-14;
Matches 119; Conservative 76; Mismatches 155; Indels 162; Gaps 19;

Qy 4 DPKIQAAMPSETISP-----MKTRSLISLLCLLSCSSWLPPLBERTESRHFNTPSKVLL 58

Db 91 DEQLAALKNFNSNSNYQIVKFKEMIQMLL-----YNNAAFLTT 129

Qy 59 DNILQIRHTPHNGLSDIYLLDDPHEALAAALIESAEHSLDLOYIWRNDISGRLLFN 118

Db 130 DNDLKI-YTDQOEKFD--LIQD-----IRNATDYIHFYIYIQNDLGRITLN 175

Qy 119 LMYLAAERGVRVRLLLDDNNTRGLDALLALDHPNIEVRLFPVLRKWRALGYLTDPP 178

Db 176 ELGKKAQGVKILYDDMGSGLL-----RKGLRFRNKGHAEAF---FP 219

Qy 179 -----RLNRRMHNSFTADNRATILGGRNIGDEYFKVGEDTVFA---DLDILATGSV 227

Db 220 SKPLINLRNMNRNHRKIVVIDGQIGVGGFNVGDEY--LGSKKFGVWRDTHLRIVGDA 277

Qy 228 VGEVSHDFRYWASHSAHNATR-----IIRSGNIGKGLQALGYNDETSRHALL 275

Db 278 VNALQURFILDMNS---QATRDHISYDDRYFPDVNSGGTIGVQIASSGPDEE----- 326

Qy 276 RYRETVEQSPLYQKIOTGRIDWQSVQTRLSIDTPAKGLDRDRRKPIAGRILQDALKQPEK 335

Db 327 -----WEQIKGYL-----KMISNA----- 343

Qy 336 SVYLVSPYFPTKSGTDALAKLVODGIDVTVLNLSQATDVAHVSHGYKVRKPLLKAGI 395

Db 344 SIVIQSPYFIPDQAFILDSIKIAALGGVDVNMIPN--KPDHPPVFWMATLKNAASLLDAGV 401

Qy 396 KLYELOPNHAPVATKDKGLTSSVTSLSHAKTFIVDGKRIFGSFNLDPRARLNTENGWV 455

Db 402 KVFHY-----DNGF-----LHSKTLVIDDEIASVGTANMDHRSFTLNFVNAF 444

Qy 456 IESPKIAEQMERTLAD-----TPPEYAYR 479

Db 445 IYDQQLAKLQKQAFIDDLAVSSSELTKARYAKR 476

RESULT 15

AI1433

cardiolipin synthase homolog lin0008 [imported] - Listeria innocua (strain Clip11262)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AI1433

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M

ok, C.; Schluetter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Vosse, H.; Wehland

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AI1433

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-504 <GLA>

A:Cross-references: UNIPROT:Q92FU5; UNIPARC:UPI00000CCOB3; GB:AL592022; PIDN:CAC95241.1

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin0008

C:Superfamily: cardiolipin synthase

Query Match 11.0%; Score 300; DB 2; Length 504;

Best Local Similarity 23.0%; Pred. No. 1e-13;

Matches 118; Conservative 76; Mismatches 210; Indels 108; Gaps 16;

Qy 14 TISPMKTRSLISLLCLLSCSSWLPPLBERTESRHFNTPSKVLLDNLQIRHTPHNGL 73

Db 90 TQVMEKAKLINAIHAIPTNNTNEKLPRLSKRI--AHLTSIEPI-----KG- 132


```
QY 365 TVLTNSLOATDVAHVSGYVKYKPKLLKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHA 424
D 361 RVMIPI--KPDHAFYVYRATTNVAGELMETGAKIF-----IYDNF-----IHA 401
QY 425 KTFIVDGKRIFIGSNLDPRLNTGNGVIESPKIAEQMERTLAD-----TTPB-Y 476
D 402 KTLVVDGEIASVGTANMDFRSLNFEVNAFIYEKKMVQKLEDAFLDILKSYQLTPELY 461
QY 477 AYR 479
D 462 AKR 464

RESULT 18
AG1387
cardiolipin synthase homolog lmo2503 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1387
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294: 849-852, 2001
A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1387
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-482 <GLA>
A;Cross-references: UNIPROT:Q8V4E3; UNIPARC:UPI000055715; GB:NC_003210; PIDN:CAD00581.1
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2503
C;Superfamily: cardiolipin synthase

Query Match 10.6%; Score 287; DB 2; Length 482;
Best Local Similarity 25.58; Pred. No. 8.1e-13;
Matches 108; Conservative 57; Mismatches 184; Indels 94; Gaps 15;

QY 74 SDIYLLDDPHEALAAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAAGVRVRL 133
D 119 NEVELFVDGHEKFDALADIEKADHILYIIPHSDELGNRLMRVLKAAEGLNVKII 178
QY 134 LDDNNTGLDLLLLALDSHPNIEVRLFPVLRKWRALGYLTDPPLNRRMHNKSFADN 193
D 179 YDAMGSRRTTKSFRTFKNGGLVRPFPKLP-----LINF-RLNRYNHRKLAIDG 230
QY 194 RATTILGGRNIGDEYFKVGEDTVFADLDILATGSVVGEVSHDFDRYWASHSAHNATRIIRS 253
D 231 DVGIGFNLGDEY--LGAKKFG-----YMRD-----TLRVH 262
QY 254 GNIGKGLQA---LGYNDETSHAL---LRYRETV--PQSPLYQKIQTG-RIDWQSVQTRL 304
D 263 GKAVYMQTRFIMDNWSASTHKIDYKARYFPFHGKGHTSMQIVSSGPDSEWQIKNGY 322
QY 305 ISDTPAKGLDRDRKPIAGRLQDALQKPEKSVYLVSPYVPFKSGTDALAKLVQGDIV 364
D 323 IK-----MINAAKTIYLSQSPYFIPDASLLLEAIKIAALSGVDV 360
QY 365 TVLTNSLOATDVAHVSGYVKYKPKLLKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHA 424
D 361 RVMIPI--KPDHAFYVYRATTNVAGELMETGAKIF-----IYDNF-----IHA 401
QY 425 KTFIVDGKRIFIGSNLDPRLNTGNGVIESPKIAEQMERTLAD-----TTPB-Y 476
D 402 KTLVVDGEIASVGTANMDFRSLNFEVNAFIYEKKMVQKLEDAFLDILKSYQLTPELY 461
QY 477 AYR 479
D 462 AKR 464
```

RESULT 19

```
G89906
hypothetical protein SAIL55 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: G89906
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Og
ma, A.; Mizutani-Ui, Y.; Hattori, M.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Tattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: G89906
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <KUR>
A;Cross-references: UNIPROT:Q99UF8; UNIPARC:UPI000000CAA52; GB:BA000018; PID:g13701116;
A;Experimental source: strain N315
C;Genetics:
A;Gene: SAIL55
C;Superfamily: cardiolipin synthase

Query Match 10.5%; Score 284.5; DB 2; Length 505;
Best Local Similarity 21.7%; Pred. No. 1.3e-12;
Matches 103; Conservative 79; Mismatches 158; Indels 135; Gaps 15;

QY 63 QIRHTPHNGLSDIYLLDDPHEALAAALIESAHSLSLDQYIWRNDISGRLLFNLMYL 122
D 143 KVDHFIDGNDLYDQVLKD-----IKNAKEYIHLEYVTFALDGLGRILHALEE 190
QY 123 AARGVVRVRLLLDDNTRGLDLLLLALDSHPNIEVRLFPVLRKWRALG-----YLT 176
D 191 KLKQGLEVKILYDDVGSK-----NVKMANFDHF-----KSLGGEVEAFPAK 232
QY 177 FP----RLNRRMHNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVGEVS 232
D 233 LPLLNFPMNRRNRKLIIVDGQGYGVGGFNIGDEYLGKGLGYWRDTHLRIQGDVADALQ 292
QY 233 HFDPRYWASHSAHNATRII-----RSGNIGKGLQALGYNDETSHALLRYRTVBQSP 285
D 293 LRFILDNWS--QAHRPQFQFVDVKYFPKKGKPLG-----NSP 326
QY 286 LYQKIQTGRIDWQSVQ---TRLISDTPAKGLDRDRKPPPIAGRLQDALQKPEKSVYLVSP 342
D 327 IQIAASGPASDWHQIYGYTKMMSA-----KKSIVYLSQP 361
QY 343 YFVPTKSGTDALAKLVQGDIDVTLTNSLQATDVAHVSGYVKYKPKLLKAGIKLYELQP 402
D 362 YFIPDNYSYNAIKIAAKSGVDVHLMIPC--KPDHPLVYWATFNSADLLSSGVKIY---- 415
QY 403 NHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIFIGSNLDPRLNTPRSGARLNTGNGVIESPKIA 462
D 416 -----TYENG-----IHSKMCLIDDEIVSVGTANMDFRSLNFEVNAFYDENLA 462
QY 463 EQMERTLADTTPYAYRVTLTKNRLQHDHPATKTYPNPEPEAKLWKRIAAKILS 517
D 463 KDL-----RVAYEHDITKSKQL-----TKESYANRPLSVKFKESIAKLVS 502
```

RESULT 20

```
H86744
cardiolipin synthase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86744
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrh
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: H86744
A;Status: preliminary
```


QY 354 LAKLVODGIDVTVLNSLOATDVAHVSGYVYKRPKLLKAGIKLYELOPNHAVPATKDKG 413
Db LLDAAARGVAVRLVLP--KKGDPVLVKEAAPFYFKELLEGGINIYRY------RG 313
QY 414 LTGSSVTSLHAKTFIVDGKRFIFGSNLDPRSARLNTMGVVIESPKIAEQMERTLADTT 473
Db 314 F-----FHGAIVIDDKLADVGTANFDKRSFRLNVEINCLLYDKEMIQVLREEL----- 362
QY 474 PEYAYRVTLDKHNLQWHDPPATRTKYTPNEPE 504
Db 363 -DYDFSIS---ERLOWMEDLANRSGFFHRTKE 388
RESULT 23
T43863
cardiolipin synthase [imported] - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43863
R:Koyama, M.
submitted to the EMBL Data Library, August 1998
A:Description: Clostridium perfringens hem operon.
A:Reference number: 222707
A:Accession: T43863
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-476 <KOY>
A:Cross-references: UNIPROT:Q9ZNC6; UNIPARC:UPI000016EODF; EMBL:AB017186; PIDN:BAA74786.
A:Experimental source: strain NCTC8237; vegetative cell
C:Genetics:
A:Gene: cIs
C:Superfamily: cardiolipin synthase
Query Match 10.0%; Score 273; DB 2; Length 476;
Best Local Similarity 24.3%; Pred. No. 8e-12;
Matches 108; Conservative 79; Mismatches 167; Indels 90; Gaps 20;
QY 41 LEERTSRHNTSKP-VLLDN-----LQIRHTPHNGL-----SDIYLLDDPHEALAA 89
Db LDDTKQKYLNSFKSHYKLDNISLUKYKDLIMNPNNDNSTVTQRNDIDLFDANSLFEEM 125
QY 90 AALIESAHSLDQYIWRNDISGRLLFNMLYLAERGVRVRLLLDD--NNTRGLD-DLL 146
Db 126 IDEINKAEKTIHMEFYIFKSDIEIGKILQALTCKAKEGVEVKLLVDSIGNSIHKKIDKL 185
QY 147 LALDSDHNIENVLPFVLKRWALGYLTDFPRLNRMHNKSFTADNRATILGRNIGDE 206
Db 186 KAAGG---DFKIFFPGFCK-----YIN--LRINYRNHRKILIDSKVAFLOGFNGIDE 233
QY 207 YFKVGEDTVFADLLDILATGVSVGVSHDFDRYWASHSAHNATRIIRSGNI--GKGLQALGY 265
Db 234 Y--LGKDK-----NIGH-----W--RDTHTKIKGLAINDLEGFLDWSY 269
QY 266 NDETSRIA--LLRYRETVEQSPLYQKIQTGRIDWOSVTRLISITPPAKGLDRDRKPPPIAG 324
Db 270 ANESDLDIDLKYPINPHSTDLPKKI-----IGAQIVSSGP-----DHTEQOIKN 314
QY 325 RLQDALQPEKSVYLVSPYFPTKSGTDLAKLVODGIDVTVLNSLOATDVAHVSGYV 384
Db 315 GYFKIINSKKNLFITQPYFVDPPEMLEALRLAALSQVDVKIMLPF---NPDKHFMGWI 370
QY 385 --KYRKPLLKAGIKLYELOPNHAVPATKDKGLTSSVTSLHAKTFIVDGKRFIFGSNLD 442
Db 371 ANSYFESLLNAGAKIY-----LYEKG-----LHAKTIVADSSICSVGTANMD 413
QY 443 PRSARLNTMGVVIESPKIAEQME 466
Db 414 IRGSFLNFESNIFIYNEAISKSME 437

RESULT 24
H82711
cardiolipin synthase XFI209 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: H82711
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82711
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <SIM>
A:Cross-references: UNIPROT:Q9PE19; UNIPARC:UPI00000C265D; GB:AE003954; GB:AE003849; NI
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir,
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XFI209
C:Superfamily: cardiolipin synthase

Query Match 9.9%; Score 270.5; DB 2; Length 467;
Best Local Similarity 27.3%; Pred. No. 1.2e-11;
Matches 117; Conservative 56; Mismatches 162; Indels 93; Gaps 18;
QY 78 LDDPPEALAAALIESAHSLSLDQYIWRNDISGRLLFNMLYLAERGVRVRLLLDD- 136
Db 114 LLDQYQSSRLLLNDIDSACERVHLLYLLMFDDAVGDAVVAALQRAAARGVSCRLLLDAV 173
QY 137 NNTRGLDLLALLDSDHNIENVLPFVLKRWALGYLTDFPRLNRMHNKSFTADNRAT 196
Db 174 GAKRGLRAYQCCCLRAH-GVEVHAMLPGLR-WRRSG-----RMDLRNHRKIAVIDNRVA 225
QY 197 ILGRNIGDEYFKVGEDTVFADLLDILATGVSVG-EVSHDFDRYWASHSAHNATRIIRSGN 255
Db 226 YIGSQNLACPDFV-----PHFPNRELVA--RVCQPAVSH----- 257
QY 256 IKGGLQALGYND---ETSRHALLRYRETVEQSPLYQKIQTGRIDWOSVQTRLISDTPAKG 312
Db 258 ----LEAVFLSDWFVETGQR--LRMSDV---PICE-----QNVAAQLVASGPAYP 299
QY 313 LDRDRKPPITAGRLQDALQPEKSVYLVSPYFPTKSGTDLAKLVODGIDV--TVLTNSL 371
Db 300 YENAR-----DAVNTMIHLARRRVLLTTPYFVDPDEATLSALRIAGTSGVDVQLILSESN 353
QY 372 QATDVAHVSGYVYKRPKLLKAGIKLYELOPNHAVPATKDKGLTSSVTSLHAKTFIVDG 431
Db 354 NQFLMAWAQEAIFY---EELLRAGIKIYALRPHF-----LHAKHLSVDD 393
QY 432 KRIFIGSNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRVTLDKHNLQW 491
Db 394 DIALFGSINLDRSFALNAEIGMICVDSRIVSRLREVESD-----YLANACQLNLVEWQ 447
QY 492 D-PATKRT 498
Db 448 QRPWERS 455

RESULT 25
B82971
cardiolipin synthase PA5394 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: B82971
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 953-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
C;Accession: B82971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <STO>
A;Cross-references: UNIPROT:Q9H7H0; UNIPARC:UPI00000C5PFD; GB:AE004952; GB:AE004091; NID:20437337
A;Experimental source: strain PA01
C;Genetics:
A;Gene: cis; PA5394
C;Superfamily: cardiolipin synthase

Query Match 9.9%; Score 268; DB 2; Length 490;
Best Local Similarity 24.5%; Pred. No. 1.9e-11;
Matches 111; Conservative 60; Mismatches 186; Indels 96; Gaps 13;

Qy 37 WLPLEERTESRHFNTSPVLNDNIQLIRHTP--HNGLSDIYLLDDPHEALAAALIE 94
Db 97 WRPWVEEALAAQVSGYKG--LKALVRMTRTPTLANN---RVRLLVNGEAFGAFMKALS 151

Qy 95 SAEHSLDLQYIWRNDISGRLLFNMLYLAERGVRVRLLLDDNNTRGLDLLLALDHPN 154
Db 152 AARQVILVQFFIVRDDLQGRLLQQLLERRAANGVEVFLLYDAIGSHALPHRYVERLQGG 211

Qy 155 IEVLFNFPVLKRWALGYLTFPRLNRRMNKSPADNRATILGGRNIGDEYFKVGEDT 214
Db 212 VQMHGFS-----TGSGMLNRF-QVFNFRHKVVVVDEGCFVGHNVGVEY--LGKRP 261

Qy 215 VFA-----DLIDLATGSVWGEVSHDF--DRYWASHSAHNATRIIRSGNIGKGLQALYNDET 269
Db 262 PLAPWRDTHMELRGPAVACLOBSFAEDWYATHSL----- 296

Qy 270 SRHALLRYRETVQSPFLQKQTGRIDWQSVQTRILISTPAKGLDRDRRKPIAGRLQDA 329
Db 297 --PPLILPPQVDSGALCQVVASGPAQAQETCSLF-----FVEM 333

Qy 330 LKQPEKSVVLPSPYVPVTKSGTDALAKLVQGDIDVTLTNSLQATDVAHVHSGYVKYRKP 389
Db 334 INAAHERVWITSFYVPDEAVMAALRLAVLRGVDRLLIPS--RPDHTVTVAASLYALE 391

Qy 390 LKAGIKLYELOPNHAVPATKDKGLTSGSVTSLSHAKTFIVDGKRFIFGSPNLDPRSARLN 449
Db 392 AIRAGVKVFRYQPGF-----LHQKVVLVDRDRTAAVGSANLDRSFLNL 434

Qy 450 TEMGVVTSFKIAEQWERTL-----ADTTPE 475
Db 435 FEVMVTVTDEGFAVEAMLEADFAESLEFTPE 467

RESULT 26
S60089
cardiolipin synthetase homolog ywiE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S60089; S55414; F70059
R;Cruz Ramos, H.; Boursier, L.; Moszer, I.; Kunst, F.; Danchin, A.; Glaser, P.
EMBO J. 14, 5984-5994, 1995
A;Title: Anaerobic transcription activation in *Bacillus subtilis*: identification of disc
A;Reference number: S60080; MUID:96112813; PMID:8846791
A;Accession: S60089
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-500 <CRU>
A;Cross-references: UNIPROT:P45860; UNIPARC:UPI0000060BCF; EMBL:Z49884; NID:9971335; PID:20437337
R;Glaser, P.; Danchin, A.
submitted to the EMBL Data Library, May 1995

A;Description: Cloning and sequencing of the *Bacillus subtilis* chromosomal region from 3
A;Reference number: S55414
A;Accession: S55414
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 432-500 <GLA>
A;Cross-references: UNIPARC:UPI000016B83F; EMBL:Z49782; NID:9853752; PIDN:CAA89861.1; PI:1
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuani, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Rieger, M.; Rivolta, C.; Roche, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. *Bacillus subtilis*.
A;Title: The complete genome sequence of the Gram-positive bacterium
A;Reference number: A69580; MUID:99044033; PMID:9384377
A;Accession: F70059
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-500 <KUN>
A;Cross-references: UNIPARC:UPI0000060BCF; GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CA
A;Experimental source: strain 168
C;Genetics:
A;Gene: ywiE
C;Superfamily: cardiolipin synthase

Query Match 9.8%; Score 266.5; DB 2; Length 500;
Best Local Similarity 22.4%; Pred. No. 2.5e-11;
Matches 107; Conservative 64; Mismatches 176; Indels 131; Gaps 15;

Qy 48 RHFTSKPVLNDNIQLIRHTPHNGLSD-----LYLLDDPHEA 85
Db 94 RMYNREK---LRKULFDEKTEPTVTLKDNQQRFTYSIRAAHMMINTKSNIVKLNGBET 150

Qy 86 LAARAALIESAEHSLDLQYIWRNDISGRLLFNMLYLAERGVRVRLLLDDNNTRGLDDL 145
Db 151 FPDIFKAMRKAESVIHIEYMFKSDMLGRGMDIMMEKARQGEVRFLYDAAGSMKLARR 210

Qy 146 LLALDHPNIEVLFNFPVLKRWALGYLTFPRLNRRMNKSPADNRATILGGRNIGD 205
Db 211 DIMRMKQAGVDIVPFSPL-----KYGFFNQ--KLNFNHRKVIIDGKTGFVGLNVGK 262

Qy 206 EYFKVGEDTVFADLLDILATGSVWGEVSHD--PDRYWASHSAHNATRIIRSGNIGKGLQALG 264
Db 263 EY-----ISRDPYIGWRD-----THLRLEGIVQTLHAIF 293

Qy 265 YND--ETSRHALLRYRE-----TVEQSPLYQKIQOTGRIDWQSVQTRILISTPAKGLDRDR 317
Db 294 MLDWEYVSNVILDOEEYNTVPVVEGGGIYQVATG-----PD 331

Qy 318 RKPPIAGRLQDALQPEKSVVLPSPYVPVTKSGTDALAKLVQGDIDVTLTNSLQATDVA 377
Db 332 MKESMSLDLYEMISSAQKSIWIATPFYFPNIESIRLTAKAAATKGEVRFVMPV--EKNDSP 389

Qy 378 AVHSGYVKYRKPRLKAGIKLYELOPNHAVPATKDKGLTSGSVTSLSHAKTFIVDGKRFIFG 437
Db 390 LTQVASSYFPPELLLEGIEVYSYQ-----KGF-----MHQKVMIDGDLASVG 432

Qy 438 SFNLDPRSARLNTMGVVI-----ESPKIAEQ--MERTLADTPE 475
Db 433 TANMDRSFQLNPEVNVFFTDAAIRLTLEAHFEEDMQESEKLSVPGFYKRGVADRTKE 490

RESULT 27
AE3539
cardiolipin synthetase (BC 2.7.8.-) [imported] - *Bruceella melitensis* (strain 16M)

C;Species: Brucella melitensis
C;Date: 01-Feb-2000 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AE3539
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AE3539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <KUR>
A;Cross-references: UNIPROT:Q8YDD8; UNIPARC:UPI00000563DD; GB:AE008918; PIDN:AAU53480.1;
A;Experimental source: strain 16M
C;Genetics:
A;Map position: 11
C;Superfamily: cardiolipin synthase
C;Keywords: transferase

Query Match 9.7%; Score 263.5; DB 2; Length 492;
Best Local Similarity 23.0%; Pred. No. 4.1e-11;
Matches 98; Conservative 69; Mismatches 180; Indels 79; Gaps 13;

QY 48 RHNTSTKPVLLDNLQIRHPTNNGLSDIYLLDPHEALARAALIESAHSLSLDQYIYW 107
DB 106 RQFGAMK-ILGDVSLYDFTSGNT-----IEMLEAGDEAYAAMLGAIGRAERSIVLETYIF 160
QY 108 RNDISGRLLFNMLYLAARGVVRVLLDDNNTRGLDLDLALDSDHPNIEVLFNPFVLRK 167
DB 161 DHDIAIGKFFADALGDVQGVGVRLVDVAGARYSPFSIVKLKKEGKGVAVFNGNIIG 220
QY 168 WRALGYLTDPPRLNRRMHNKSFADNRATILGGRNIGDEYFK-VGEDTVFADLDILATGS 226
DB 221 LR-----LPEANLRTHRWMLVDGQTAFVGGNNIRAGFVRAIGADAVAFDFHKEGP 273
QY 227 VGVGVSHDFRYWASHSAHNATRIIRSGNIGKGLQALGYNDTSRHALRYRETVEOSPL 286
DB 274 AIADLFHIASEDNR-----FATGELLTG-----EAWSTAPP 304
QY 287 YQKIQGTGRIDWQSVOTRLSDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVLYSPVFP 346
DB 305 ENPPGTGTL-----VAVWGGPGDKLETNHR-----MMGAFSIAQOHILIMTPYLLP 352
QY 347 TKSQTDALAKLVQDGDVTVL---TNSIQATDVAHVSGVYKPKLLKAGIKLYELQPN 403
DB 353 DRELISALVTAARGVSDIVVPGVNNKLVD-RAMPAQF-----DQLRDGCRIRW---- 403
QY 404 HAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFGFSNLDPRSARLNTMGVIESPKIAE 463
DB 404 -----AGGAFFN-HSKLMTIDGMSYVGVSSNIDPRSLNFEVLDLEILDRDVAR 450
QY 464 QMERTL 469
DB 451 QVEERI 456

RESULT 28
D83103
probable phospholipase PA4339 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: D83103
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lm,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <STO>

A;Cross-references: UNIPROT:Q9HW62; UNIPARC:UPI000000CSCCA; GB:AE004850; GB:AE004091; NI
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4339
C;Superfamily: cardiolipin synthase

Query Match 9.6%; Score 261; DB 2; Length 359;
Best Local Similarity 23.8%; Pred. No. 3.8e-11;
Matches 93; Conservative 57; Mismatches 148; Indels 92; Gaps 10;

QY 93 IEAHSLSLDQYIWRNDISGRLLFNMLYLAARGVVRVLLDDNNTRGLDLDLALDSDH 152
DB 5 IEAQRSIELELYVEDHGHCAELFLVALLDARRRGVAVRCLFDGFGCLGLGSAWIOQLRE 64
QY 153 PNIEVLFNPFVLRKWRALG---YLTDFFPRLNRRMHNKSFADNRATILGGRNIGDEYFK 209
DB 65 AGGELALYNPL---RWKLTGGNLY-----RDHRKLLLVGRLGYGVGAGITDFWE 112
QY 210 -VGEDTVFADLDILATGSVVGVSHDFDRYWASHSAHNA-----TRIIRSGNIGKG 259
DB 113 PVSIVSAWREVMVMDGPPVADWALFERQWLACLEKAWKPREGMTLRLPPQGAARG 172
QY 260 LQALGYNDTSRHALRYRETVEOSPLYQIKIQTGRIDWQSVOTRLSDTPAKGLDRDRK 319
DB 173 LGRVAYAD-----ARQHRDILQSLVRALNGSRRR 201
QY 320 PPIAGRLQDALKQPEKSVLYSPVFPVTPKSGTDALAKLVODGIDVTVLTNSLOATDVAV 379
DB 202 -----IWLATYFLEPTWKVRALKAQGVREVRLLAG-RLTDHAPV 243
QY 380 HSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFG 439
DB 244 RYAGQRYVPRLLRAGVRIHEYQPRF-----LHLKMWVD-DWVSVSGC 285
QY 440 NLDPRARLNTMGVIESPKIAEQMERTL 469
DB 286 NFDHNLRLFNLDANLEALDPDFTNEAASL 315

RESULT 29
E97307
probable cardiolipin synthase (phospholipase D family) [imported] - Clostridium acetob
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97307
R;Nolling, J.; Breton, G.; Omselchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-510 <KUR>
A;Cross-references: UNIPROT:Q97E04; UNIPARC:UPI000000D756C; GB:AE001437; PIDN:AAK81248.1
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3316
C;Superfamily: cardiolipin synthase

Query Match 9.6%; Score 260; DB 2; Length 510;
Best Local Similarity 20.8%; Pred. No. 7.6e-11;
Matches 103; Conservative 79; Mismatches 177; Indels 136; Gaps 16;

QY 36 SMLPPLERTESRHFNSTKPVLLDNLQIRHPTNNGLSDIYLLDDPHEALARAALIES 95
DB 138 SLPPPY-KYTKAEYFSTGETFNEELQLK-----X 167
QY 96 AEHSLSLDQYIWRNDISGRLLFNMLYLAARGVVRVLLDDNNTRGLDLDLALDSDHPN- 154
DB 168 AKQYIFLEYFITKEGVMMNSVLQILDRKVOEGVEVVIYDD-----IGCMFTLPG 218
QY 155 -----IEVRLFPFVLRKWRALGYLTDPPRLNRRMHNKSFADNRATILGGRNIGD 205

```
Db 219 YHKELEKIGIKCCVFNPLI-----PMWSF-KFNNDHRKRIAVIDGLVGTGINUSD 269
Qy 206 EYF-KVGSDTVFADLDILATGSVVGESHDFDRYWASHSAHNATRIIRSGNIGKGLQALG 264
Db 270 EYINKYQYGYWKDTAIFKEGKAANLWSVFLSMWS-----FLR-GIDEDFNFTFK 318
Qy 265 YNDETSRHALLRYRETVEQSPLYQKIQTRIDWQSVQTRLISDTPAKGLDRDRKPPITAG 324
Db 319 KNIKEFSRSREGYVQPFADSEL-----DGEVPGEIYYNMLIS----- 355
Qy 325 RLQDALKOPEKSVYLVSPYFPTKSGTDLAKLVQDGDVTLVTLNSLQATDVAAVHSGYV 384
Db 356 -----KATKVYITTPYLIVIGNEMVTALSAAGGVDRRIITPHI---PKKIVHSVTK 406
Qy 385 KYRPELLKAGIKLYELOPNHAVPATKOKGLTSSVTSLSHAKTFIVDGKRIFIGSNLDPR 444
Db 407 SYKVLIESGVKIYEMPGF-----IHSKTYVCDNEYGVGVSINMDFR 449
Qy 445 SARLNTMGVWIESPKIAEQMERTLADTTPEYAYRVTLTKHNLQWHDPATRKYYPNEPE 504
Db 450 SLYLHFECGVWMYKNTNTYDIKKDFMDTL-DKSKEITLEEINKVKWY----- 495
Qy 505 AKLWK--RIAAKIL 516
Db 496 STLWRVLRVAPLM 510

RESULT 30
AH2333
cardiolipin synthase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2333
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2333
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <KUR>
A:Cross-references: UNIPROT:Q8YPH1; UNIPARC:UPI00000CEA63; GB:BA000019; PIDN:BAB75922.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4223
C:Superfamily: cardiolipin synthase

Query Match 9.4%; Score 256; DB 2; Length 480;
Best Local Similarity 24.3%; Pred. No. 1.4e-10;
Matches 120; Conservative 69; Mismatches 181; Indels 124; Gaps 18;

Qy 35 SSWLPPLLEERTSRHFNTPSKVLLDNILQIRHTPHNGLSDIYLLDDPHEALAAALIE 94
Db 98 SAKUEPLQ-----LLAEAFIGIFPTSGNNA-----KLLINGQOQTYAAMLISA 140
Qy 95 SAEHSLDIQYIWRNDISGRLLFNLMYLAABERGVRVRLLLDDNNTRGLDLLLLALDSHPN 154
Db 141 SANSYILLOSYIVVDDKAGNEFKDALIAKAGQIRVYLIYDEIGSNKISRLYVKSLOKYD 200
Qy 155 IEVRLFNPFVLRKWRALGYLDFPPLNRRMKNKSTADNRATILGGRNIGDEYFKVGED- 213
Db 201 IQVSAPHTTRGRGNRF-----QLNFRNHRKILVDGRTAFIGGLNISDEY--LGKNP 250
Qy 214 --TVFADLDILATGSVGEVSHDF--DRYWASHSAHNATRIIRSGNIGKGLQALGYND 269
Db 251 RLSPWRDTHMLLEGTPTVOSLOGCFLODWYAT----- 282
Qy 270 SRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLISDT-PAKGLDRDRKPPITAGRL-- 326
Db 283 -----RQVIDVNWQVP-----NWESDYTALVFPTGPA-----DKLK---ACKLFF 320
```

```
Qy 327 QDALKOPEKSVYLVSPYFPTKSGTDLAKLVQDGDVTLVTLNSLQATDVAAVHSGYVKY 386
Db 321 VSAINQOQTRLIWATPYFVDDSTLTALKLAALRGVDVRIILPN--RPDHLVLVLCFSFSY 378
Qy 387 RKPLLKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIFIGSENLDPESA 446
Db 379 YTEMKATNIKLYRYK--HGF-----MHQKVILIDKEWAGVGTVNLDRNSP 421
Qy 447 RLNTE-MGVWIESPKIAEQMERTLADTTPEYAYRVTLTKHNLQWHDPATRKYYPNEPEA 505
Db 422 FLNFEVMGFVANSQFVKSVEMQLQADL--KAALAVDFS DYER-----K 462
Qy 506 KLVWRIAANKILSL 519
Db 463 YLWFKLAVRISLL 476

RESULT 31
C82171
cardiolipin synthase VC1670 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82171
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: C82171
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-484 <HRI>
A:Cross-references: UNIPROT:Q9KRH2; UNIPARC:UPI00000C307E; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1670
A:Map position: 1
C:Superfamily: cardiolipin synthase

Query Match 9.0%; Score 243.5; DB 2; Length 484;
Best Local Similarity 25.0%; Pred. No. 1.1e-09;
Matches 104; Conservative 59; Mismatches 142; Indels 111; Gaps 18;

Qy 71 NGLSDIYLLDDPHEALAAALIESAHSGLDQYIWRNDISGRLLFNLMYLAABERGVRV 130
Db 119 NTLG---LLNSPNEILHAIIDDIERIAQFQIRMFYIWHPGGLADAVASAVIQASKRGVNV 175
Qy 131 RLILDDNNT-----RGLDLLLLALDSHPNIEVRLFNPFVLRKWRALGYLTD 176
Db 176 KLLLDLSAGSPRFFSPPEKMMRDAGI-EVVOALEVSP-----WRIF----- 215
Qy 177 PPLNRNMHNKSFADNRATILGGRNIGD-EYFK-----VGEDTVFADLDILATGSVVG 231
Db 216 LRRDLQHRKILVIDDEIATYGSNNWVDPAFYKQAGVGO---WIDIMVRTGTNNVL 272
Qy 232 S--HDFDRYWASHSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQK 289
Db 273 SAHCWD--MEFET-----GSRMLPKNPE-----CRLEPNQOHPHI-QV 308
Qy 290 IQTRIDWQSVQTRLISDTPAKGLDRDRKPPITAGRLQDALKOPEKSVYLVSPYFVPKFS 349
Db 309 VPSG-----PGMPENLISQV-----LTLAINQANRSVCITTPYFVPSAD 347
Qy 350 GTDALAKLVQDGDVTLVTLNSLQATDVAAVHSGYVKYRKPLLKAGIKLYELOPNHAVPAT 409
Db 348 LLATLKMTAQRGKIVDIIP--KNDLSLMVOWASAFYGELLEAGVQIHEF----- 396
Qy 410 KDKGLTSSVTSLSHAKTFIVDGKRIFIGSNLDPRSAHLNTEMGVIESPKIASQM 465
Db 397 -DGLG-----LHTKSVVIDQQFCLVGTVNLDMRSLWNLNFEFLTAVDDLEFTQOM 444
```


RESULT 32

AG0266
cardiolipin synthetase (EC 2.7.8.-) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0266
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0266
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-486 <KUR>
A:Cross-references: UNIPROT:Q8ZEI2; UNIPARC:UPI00000CDBB3; GB:AL590842; PIDN:CAC90995.1;
C:Genetics: cis
A:Gene: cIs
C:Superfamily: cardiolipin synthase
C:Keywords: transferase

Query Match 8.9%; Score 241.5; DB 2; Length 486;
Best Local Similarity 22.5%; Pred. No. 1.5e-09;
Matches 112; Conservative 81; Mismatches 180; Indels 125; Gaps 23;

QY 33 SCSSWLPLEBTERSRH-ENTSKPVLDDNLIQIRHTPHN-NGL--SDIYLLDDPHEALAA 88
Db 77 STAKWLSLK---ECQHFANSNEVASPLFCQRCRGGVKGNGQLQLTTTDDTLKA 133
QY 89 RAALIESAEHSLDQYIWR-NDISGRLLFNLMYLAERGVVRVRLDDNNTR----- 140
Db 134 LVRLDIELARHNIEVFIWQPGVLVDQVAESLM-AAARRGVHCELLDLSAGSKQFFRSPY 192
QY 141 -----GLDLLLLALSDHPNIEVRLFPVLRKWRALGYLTDPPRLNRMRHNSFTADN 193
Db 193 PAMMRNAGIEV-----EALKVNVFRMFL-----RRMDLRQHRKIVLIDN 232
QY 194 RATILGGRNIGDEYFKVGEDTVFADLDILA-----TGSVVGEVSHDFRYWASHAHNAT 248
Db 233 YVAYTGSMNMVDPFRFFQDAGVGQWIDMARMGEFVATTIGV-----YACDWEIETGK 286
QY 249 RII---RSGNIGKGLQALGYNDETSRHALLRYRTVEQSPLYQKIQTGRIDWQSVQTRLI 305
Db 287 RILPPPPDANI-----MPFEETG-HTI---QVIASGQFPE-----EMIHCALL 327
QY 306 SDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTGDALAKLVODGIDVT 365
Db 328 T-----AVYAAREQLIMTTPYFVPSDDLHAICTAAQRGVDVS 365
QY 366 VLNTSLQATDVAAVHSGYVKYKRPKLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHA 425
Db 366 IIVP--RENDMMVRWASRAFFTELLNAGVKIYQF-----EGGL-----LHSK 406
QY 426 TFIVDGKRIFIGSNLDPRSLNTEMGVVIESPKEIAQOMERTLADTPPEYAVRVTL--- 482
Db 407 SVLVGQLSLVGTVNLDMRSLNFEITLVIDD-----DGFADLAQVDDYIARSALLDG 462
QY 483 DKHN-LQWHDPAIRTKTY 499
Db 463 ERWKRPLMHRVTERLFY 480

RESULT 33

S55419
cardiolipin synthetase homolog ywJ - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S55419; C70060
R:Glaser, P.; Danchin, A.
A:Description: Cloning and sequencing of the Bacillus subtilis chromosomal region from 3

A:Reference number: S55414
A:Accession: S55419
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <GUA>

A:Cross-references: UNIPROT:P45865; UNIPARC:UPI0000060BCA; EMBL:249782; NID:G853752; P1
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bert
C.; Bron, S.; Broutillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C
A.; Ehrlich, S.D.; Emerson, P.T.; Etian, K.D.; Errington, J.; Fabret, C.; Ferrari, E
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tetrsta, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C70060
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-398 <KUN>

A:Cross-references: UNIPARC:UPI0000060BCA; GB:Z99123; GB:AL009126; NID:G2636240; PIDN:C
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywJ
C:Superfamily: cardiolipin synthase

Query Match 8.6%; Score 235; DB 2; Length 398;
Best Local Similarity 20.7%; Pred. No. 3.3e-09;
Matches 93; Conservative 71; Mismatches 178; Indels 108; Gaps 12;

QY 74 SDIYLLDDPHEALAAALIESAEHSLDLOYYTWRNDISGRLLFNLMYLAERGVVRVRL 133
Db 43 SDIELHCGADIVRMWMDIRQAASSVHMFFIMKDEVSHNMVTLTKTKAQAGSVYLL 102
QY 134 LDNNTRTGLDLLLLALSDHPNIEVRLFN-----PFVLRKWRALGYLTDPPRLNRMRHNSF 189
Db 103 LDWAGCRAIKKTALQTMKNAGVHVVMNRPRFPFF-----FFHMQRNHRKIT 150
QY 190 TADNRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVEVSHDFRYWASHAHNAT 249
Db 151 VIDGKIYIGGFNIAIBEY--LGKKAKFGNWEYHL-RMIGEGVHDLQTLFASDLKRNT-- 205
QY 250 IIRSGNIGKGLQALGYNDETSRHALLRYRTVEQSPLYQKIQTGRI-----DWQSVQT 302
Db 206 -----GIE-LG-----SDVWPKLQOGTISHKIYATDGYSLN 236
QY 303 RLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTGDALAKLVODGI 362
Db 237 IYLAN-----IAQAKNRLTVCTPYIPSPKPLQEQALINARKNGV 274
QY 363 DVTVLNTSLQATDVAAVHSGYVKYKRPKLKAGIKLYELOPNHAVPATKDKGLTGSSVTS 422
Db 275 SVRIIVP--MKSHPLVREAAFTYISELLDAGCLYIYYQGF-----Y 315
QY 423 HAKTFIVDGKRIFIGSNLDPRSLNTEMGVVIESPKEIAQOMERTLADTPPEYAVRVTL 482
Db 316 HVKALLIIDDHLSIIGTANFDKSLFLNEEVNVEIDDEAFTSEYVATIEEDM----- 366
QY 483 DKHNRLQWHDPAIRTKTYPNEPEAKLWKRIA 512
Db 367 -KKSELLTMEDFSRTFRQPAEWLGRALS 395

RESULT 34

AI0598
probable phospholipase ybho [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi

Qy	264	GYNDETSRHALLRYRETVEGSPLYQKI--OTGRIDWQSVQTRLLISDTPAKGLDRDRKPP	321
Db	283	-----SVRDPKOKLEREDSYFYVPEEGNIDIQ-----IVANGPKS--DNKTLRTG	326
Qy	322	IAGRLQDALQPKSVVLYSPVPYFPTKSGTGDALAKLVQDGDIVTVLNSLQATDVAAVHS	381
Db	327	FIKWRMDA-----BDYIWLQSPYIIPDDSMITALVAANSQVDVIRIMIPNW--PDHPFIFR	380
Qy	382	GVYKYRKPLLKAGIKLYELQPNHNAVPATKDKGLTGSSVTSHTAKTFIVDGKRIFIGSFNL	441
Db	381	ATOYYANYLHKHGVKIYNYNGF-----IHSKTLVMDCKGLGVFGTTNQ	423
Qy	442	DPRSARLNTMGVVVIESPKIAEQMERT	468
Db	424	DIRSYELNFEISAFYDETVAKEMSR	450

RESULT 36

AC2882

cardiolipin synthase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AC2882

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y. M.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutuyavin, T.; Levy, R.; Li, M.; X Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gorder, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AC2882

A:Molecule type: DNA

A:Residues: 1-492 <KUR>

A:Cross-references: UNIPROT:Q8UCJ8; UNIPARC:UPI000000D1F08; GB:AE008688; PIDN:AA

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: C1s

A:Map position: circular chromosome

C:Superfamily: cardiolipin synthase

Query Match	8.3%	Score	224.5;	DB 2;	Length	492;			
Best Local Similarity	23.0%;	Pred.	No. 2.6e-08;						
Matches	96;	Conservative	56;	Mismatches	190;	Indels	75;	Gaps	11;

Qy	48	RHFNWTSKPVLLNLIQIRHT--PHNNGUSDIYLLDDPHEALAAARALTESAEHSLDLOYVI	106
Db	106	RRFEAMK--TLGDRV--TRHALTTGNGIEPLVSGD---VAYAAMLEAIGCAKRSIILEYI	159
Qy	107	WRNDISGRLLFNLMYLAERGVRVRLDDNNNTGRLDGLDLLALDSDHNEVRLNPFVLR	166
Db	160	FDNDRIGARFVAALERAKURGVEVRVRLDVGARYSVPSILPTLRDKDIYVADVFNQVIM	219

```
Qy 167 KWRALGYLTDPRNLRMHKSFADNRATILGGRNIGDEY - FKVGEDTVFADLDILATG 225
Db 220 GLR-----LPYANLRTHRKILVVDGRIAPSGMNIREGFTLEFGESQSHDTFKITG 272
Qy 226 SVGGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRTVEQSP 285
Db 273 PVVSDFFSIAAEDWRFTTGE-----VLDAP 297
Qy 286 LYQKIOTGRIDWQSVOTRLISDTPAKGLDRDRRKPPIAGRLQDALKOPEKSVLVSPYFV 345
Db 298 WMDIAPDGPVGSQIVARVCSSGPDKSIETSHK-----MLMGAFVSARSILIMSPYFL 351
Qy 346 PTKSGTDALAKLVODGIDVTVLNLSQATDVAAVHSGYVKYRKPPLKAGIKLYELQPNHA 405
Db 352 PDRELISALITAAARGCVVDIIVP--KSNNLVLVDRAWTAQFDQMLKNYCIWR----- 403
Qy 406 VPATKDKGLTGSSVTSIHKATFIVDGKRIFGSNFLDPRSRARLNTMGVVIESPkia 462
Db 404 -----ATGA-----FNHSKLLVIDGRWSYIGSSNLDPRSLRNLNFIDLEVMDEEFA 449

RESULT 37
B97658
cardiolipin synthase (PA5394) [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97658
R;Goodner, B.; Winkler, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Hinkle, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <KUR>
A;Cross-references: UNIPROT:Q8UCJ6; UNIPARC:UPI00000D1F08; GB:AE007869; PIDN:AAK88219.1;
C;Genetics:
A;Gene: AGR C 4515
A;Map position: circular chromosome
C;Superfamily: cardiolipin synthase

Query Match 8.3%; Score 224.5; DB 2; Length 492;
Best Local Similarity 23.0%; Pred. No. 2.6e-08;
Matches 96; Conservative 56; Mismatches 190; Indels 75; Gaps 11;

Qy 48 RHFNTPKPVLLDNILQIRHT - PHNNGLSDIYLLDDPHEALAAARALIESAHSLSLDQYII 106
Db 106 RRFEAMK - TLGDRV - -TRHALTTGNGIEPLVSGD--VAYAAMLEAIGEAKRSIILETYI 159
Qy 107 WRNDISGRLLFNLMYLAABERGVRVRLDDNNTRGLDLDLALDSDHNPNIENVRLFPFVLR 166
Db 160 FDNDRIGARVAALERAKLGEVRVRLDIDAVGARYSVPSLPTLRDKDIDVADVFNGNVM 219
Qy 167 KWRALGYLTDPRNLRMHKSFADNRATILGGRNIGDEY - FKVGEDTVFADLDILATG 225
Db 220 GLR-----LPYANLRTHRKILVVDGRIAPSGMNIREGFTLEFGESQSHDTFKITG 272
Qy 226 SVGGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRTVEQSP 285
Db 273 PVVSDFFSIAAEDWRFTTGE-----VLDAP 297
Qy 286 LYQKIOTGRIDWQSVOTRLISDTPAKGLDRDRRKPPIAGRLQDALKOPEKSVLVSPYFV 345
Db 298 WMDIAPDGPVGSQIVARVCSSGPDKSIETSHK-----MLMGAFVSARSILIMSPYFL 351
Qy 346 PTKSGTDALAKLVODGIDVTVLNLSQATDVAAVHSGYVKYRKPPLKAGIKLYELQPNHA 405
Db 352 PDRELISALITAAARGCVVDIIVP--KSNNLVLVDRAWTAQFDQMLKNYCIWR----- 403
Qy 406 VPATKDKGLTGSSVTSIHKATFIVDGKRIFGSNFLDPRSRARLNTMGVVIESPkia 462
Db 404 -----ATGA-----FNHSKLLVIDGRWSYIGSSNLDPRSLRNLNFIDLEVMDEEFA 449
```

RESULT 38

```
B83377
probable phospholipase PA2155 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: B83377
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.B.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: B83377
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-401 <STO>
A;Cross-references: UNIPROT:Q911W0; UNIPARC:UPI00000C55D9; GB:AE004642; GB:AE004091; NI
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2155
C;Superfamily: cardiolipin synthase

Query Match 8.2%; Score 224; DB 2; Length 401;
Best Local Similarity 22.8%; Pred. No. 2.1e-08;
Matches 98; Conservative 59; Mismatches 184; Indels 88; Gaps 11;

Qy 68 PHNNGLSDIYLLDDPHEALAAARALIESAHSLSLDQYIIWRNDISGRLLFNLMYLAABERG 127
Db 4 PWRDGG-NRVELLNGEYFPRLPQCIAEARREILLETFFIFEDEVGRQLQEALSAABERG 62
Qy 128 VYVRLLLDDNNTRGLDLDLALDSDHNPNIENVRLNPPFVLRKWRALGYLTDPRNLRMHK 187
Db 63 VEQVTVDGVTASLSPDYLRATASGRVRLHFDPP-----KPRLLGMRTN---LFRRLLRHK 115
Qy 188 SFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVGGEVSHDFRYWASHAHNA 247
Db 116 LVVIDRRQAFVGGINYGEDH----- 135
Qy 248 TRIIRSGNIGK---GLQALGYNDTSRHALLRYRTVEQSPLYQKIOTGRIDWQSVOTRL 304
Db 136 --LVRRGNMAKQDYAVRVEGPPVVRDIRQACLALLEDADYVPLRPSGAG---QPARVRL 189
Qy 305 ISDTPAKGLDRDRRKPPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDALAKLVQDQIDV 364
Db 190 VI-----RNDQSSDDIERYLQAIRQARRRLLIANAYFFPGYRLRLRLRDAARRGRV 243
Qy 365 TVLTNSLQATDVAAVHSGYV-KYRKPLLKAGIKLYEL--QPNHAVPATKDKGLTGSSVTS 421
Db 244 DLVLQGMPPMLVRLCSRLLYDY--LLREGVRIHYCQRP----- 281
Qy 422 LHAKTIFVQKRIFGSNFLDPRSRARLNTMGVVIESPkiaEQMERTLADTTPEYAYRVT 481
Db 282 LHGKVAVIDDDSTIGSSNLDPLSLNLEANLVRDVAFNGLYQHLRELARRHCRIS 341
Qy 482 LDKHNEFLOW 490
Db 342 RRHARRGYW 350
```

RESULT 39

```
E64815
Ybho protein - Escherichia coli (strain K-12)
N;Alternate names: protein b0789
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: E64815
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
```

A:Accession: E64815
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-413 <BLAT>
A:Cross-references: UNIPROT:P75771; UNIPARC:UPI000013A35B; GB:AE000181; GB:U00096; NID:9
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ybhO
C:Superfamily: cardiolipin synthase

Query Match 8.1%; Score 221; DB 2; Length 413;
Best Local Similarity 21.9%; Pred. No. 3.5e-08;
Matches 89; Conservative 53; Mismatches 171; Indels 94; Gaps 10;

Qy 76 IYLLDDPHEALAAARALIESAHSLLDLOYYIWRNDISGRLLFNLMYLAARGVRRLLD 135
Db 11 IQLLENGQYYPAVFAKGEAQERIIETFIWFEDDVGKQLHAALLAAQGVKAEVLDD 70

Qy 136 DNNTRGLDDLLALDSDHNEIVRNFNPFVLRKRALGYLTDFPRLNRRMHNKSTADNRA 195

Db 71 GYGSPLSDSEFVNELTAAGVVFRIYDP-----RPLFGMRTN---VFRMRHKIIVVIDARI 123

Qy 196 TILGGRNIGDVEYFKVGEDTVFADLDILATGSGVGEVSHDFD-----RYWASHSA 244

Db 124 AFIGGLNYSAEHMSYSGPEAKQDYAVRLEGPIVEDIL-QFELENLPGQSAARRWRRH-- 180

Qy 245 HNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYOKIOTGRIDWQSVQTRL 304

Db 181 HKAEENRQPG-----EAQVLLVWR----- 199

Qy 305 ISDTPAKGLDRDRKPPPIAGRLQDALQPEKSVVLVSPYFVPTKSGTDLAKLVQDGDIV 364

Db 200 -----DNEEHRDDIERHYLKMLTQARREVITIANAYFFPGYRFLHALRKAARRGVRI 250

Qy 365 TVLTNSLQATDVAAVHSGYVKYRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLSHA 424

Db 251 KLIIQG--EPDMPIVRVGARLLYNNLYKGVQVFEYRRR-----PLHG 291

Qy 425 KTFIVDGKRIFIGSFNLDPRSARLNTENGVVIESPKIAEQMERTLAD 471

Db 292 KVALMDDHWATVGSNLDPLSLNLEANVIIHD-----RHFNQTLRD 334

RESULT 40
D85587
probable synthetase ybhO [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85587
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85587
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <STO>
A:Cross-references: UNIPROT:P75771; UNIPARC:UPI000013A35B; GB:AE0005174; NID:gi2513778; E
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ybhO
C:Superfamily: cardiolipin synthase

Query Match 8.1%; Score 221; DB 2; Length 413;
Best Local Similarity 21.9%; Pred. No. 3.5e-08;
Matches 89; Conservative 53; Mismatches 171; Indels 94; Gaps 10;

Qy 76 IYLLDDPHEALAAARALIESAHSLLDLOYYIWRNDISGRLLFNLMYLAARGVRRLLD 135
Db 11 IQLLENGQYYPAVFAKGEAQERIIETFIWFEDDVGKQLHAALLAAQGVKAEVLDD 70

Qy 136 DNNTRGLDDLLALDSDHNEIVRNFNPFVLRKRALGYLTDFPRLNRRMHNKSTADNRA 195

Search completed: May 2, 2006, 06:14:03
Job time : 45 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:06:00 ; Search time 187 Seconds
(without alignments)
1233.550 Million cell updates/sec

Title: US-10-665-990A-14

Perfect score: 2720

Sequence: 1 MHTDPKIQAMSETISPMKT.....KLWKRIAAKTLISLLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*
- 9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2712	99.7	525	6	ABP79466
2	2643	97.2	525	5	ABB78070
3	2606.5	95.8	507	3	AAy75751
4	2586	95.1	508	3	AAy75753
5	2573	94.6	508	3	AAy75752
6	831.5	30.6	543	8	ADL04719
7	790	29.0	564	7	ABO75163
8	632	23.2	502	2	AAW55452
9	632	23.2	502	2	AAW55550
10	590	21.7	428	2	AAW98378
11	493.5	18.1	271	6	ADA34272
12	478.5	17.6	275	6	ADA34304
13	419	15.4	2519	7	ABM88218
14	409	15.0	321	6	ADA36398
15	308	11.3	513	6	ABM71816
16	296	10.9	486	7	ADG30702
17	294.5	10.8	504	5	ABB49813
18	291	10.7	495	5	ABP39428
19	291	10.7	495	8	AGS04559
20	287	10.6	482	5	ABB49993
21	286.5	10.5	493	5	ABB32481
22	286.5	10.5	493	7	ADC21336
23	286.5	10.5	493	7	ADC25022
24	286.5	10.5	493	7	ADD52524

25	286.5	10.5	502	5	ABB32493	Abb32493 Staphyloc
26	286.5	10.5	502	7	ADC21402	Adc21402 Staphyloc
27	286.5	10.5	502	7	ADC25088	Adc25088 S. aureus
28	286.5	10.5	502	7	ADD52590	Add52590 Staphyloc
29	286.5	10.5	510	6	ABM71594	Abm71594 Staphyloc
30	282.5	10.4	480	5	ABB54276	Abb54276 Lactococc
31	275	10.1	420	9	ABM97264	Abm97264 M. xanthu
32	274.5	10.1	206	2	AAW55249	Aaw55249 H. pylori
33	274.5	10.1	490	6	ABM69020	Abm69020 Photorhab
34	268.5	9.9	584	7	ABO75565	AbO75565 Pseudomon
35	268	9.9	579	7	ABO68379	AbO68379 Pseudomon
36	259.5	9.5	494	5	ABP39331	Abp39331 Staphyloc
37	259.5	9.5	494	8	ADS05084	AdS05084 Staphyloc
38	252	9.3	166	6	ADA36341	Ada36341 Acinetoba
39	252	9.3	422	4	ABM76548	Abm76548 Corynebac
40	252	9.3	500	4	AAG92760	Ag92760 C glutami
41	252	9.3	500	5	AAG80237	Ag80237 C. glutam
42	246	9.0	508	7	ADF04432	Adf04432 Bacterial
43	241	8.9	409	7	ADH85779	Adh85779 Enterococ
44	232	8.5	442	4	AAG81953	Ag81953 S. epider
45	231.5	8.5	386	7	ADC94461	Adc94461 E. faeciu

ALIGNMENTS

RESULT 1

ABP79466

ID ABP79466 standard; protein; 525 AA.

XX AC ABP79466;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 5462.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Masignani V, Monaci E;

XX DR WPI; 2003-058415/05.

XX DR N-PSDB; ABZ40436.

XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection.

XX PS Disclosure; Page 587; 815pp; English.

XX CC The present invention relates to proteins from Neisseria gonorrhoeae. Also disclosed are the nucleic acid molecules encoding the proteins and antibodies that specifically bind to the proteins. The composition comprising the protein, nucleic acid or antibody is useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection, this may be in the form of a vaccine or gene therapy.

XX CC Sequences given in records ABP76736-ABP81046 represent nucleic acid molecules of the invention

XX SQ Sequence 525 AA;

Query Match 99.7%; Score 2712; DB 6; Length 525;

Best Local Similarity 99.6%; Pred. No. 9.9e-240;

Matches	523;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCCSSMLPPLLEERTESRHFTSKPVL	LDN	60					
Db	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCCSSMLPPLLEERTESRHFTSKPVL	LDN	60					
Qy	61	ILQIRHTPHNNGLSDIYLLDDPHEALAAARAALIESAHSGLQYIWNNDISGRLL	FNLM	120					
Db	61	ILQIRHTPHNNGLSDIYLLDDPHEAFARAALAIESAHSGLQYIWNNDISGRLL	FNLM	120					
Qy	121	YLAAGRVRRVRLLLDDNNTRGLLDLLLALDSHPNIEVRLNFPFLVKWRALGYLT	DPRL	180					
Db	121	YLAAGRVRRVRLLLDDNNTRGLLDLLLALDSHPNIEVRLNFPFLVKWRALGYT	DFPR	180					
Qy	181	NRRMHKNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVVGESH	DFR	240					
Db	181	NRRMHKNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVVGESH	DFR	240					
Qy	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVQEOSP	LYOKI	300					
Db	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVQEOSP	LYOKI	300					
Qy	301	QTRLISDTPAKGLDRDRKPPIAGRLQDALQKPEKSVYLVSPYFVPTKSGT	DALAKL	360					
Db	301	QTRLISDTPAKGLDRDRKPPIAGRLQDALQKPEKSVYLVSPYFVPTKSGT	DALAKL	360					
Qy	361	GIDVTVLNTSLQATDVAAVHSGYVKYRPLLKAGIKLYELQPNHAVP	ATKDKGLTGSSVT	420					
Db	361	GIDVTVLNTSLQATDVAAVHSGYVKYRPLLKAGIKLYELQPNHAVP	ATKDKGLTGSSVT	420					
Qy	421	SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESP	KIAEQMERTLADTTPEY	480					
Db	421	SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESP	KIAEQMERTLADTTPEY	480					
Qy	481	TLDKHNLQWHDPA	TRKTYPNPEAKUWKRIA	KILSLPIEGLL	525				
Db	481	TLDKHNLQWHDPA	TRKTYPNPEAKUWKRIA	KILSLPIEGLL	525				

RESULT 2
ABB78070
ID ABB78070 standard; protein: 525 AA.

XX	
DT	29-AUG-2003 (revised)
DT	05-NOV-2002 (first entry)

XX Amino acid sequence of p55 polypeptide.

AA p177; p88; p64; p55; p46; p33; p30; p27; p24; p21; p18; p15; p12; p9; p6; p3; p0
KW p177; p88; p64; p55; p46; p33; p30; p27; p24; p21; p18; p15; p12; p9; p6; p3; p0

OS *Neisseria gonorrhoeae*.

PN WO200260936-A2.

08-AUG-2002.

31-JAN-2002: 2002WO-US002881.

31-JAN-2001: 2001US-0266070P.

PK 08-AUG-2001; 2001US-0310338E.
PB 23-OCT-2001; 2001US-0344452P.

AA
PA (IOWA) UNIV IOWA RES FOUND.

PA (APIC/) APICELLA M A.
PA (APIC/) UNIV CALIFORNIA.

PA (GIBS//) GIBSON B W.

PA (BROW//) BROWN E.

PI Apicella MA, Edward

Apicella MA, Edwards JL, Gibson BW, Scheffler K, Brown E;

XX	WFI; 2002-619227/66.
DR	N-PSDB; ABQ78301.
DR	
XX	
PT	New polypeptide comprising p177, p88, p64, p55 or p46 from <i>Neisseria</i>
PT	gonorrhoeae, useful for preventing, or protecting a female patient
PT	against, <i>N. gonorrhoeae</i> colonization or infection.
XX	
XX	
PS	Claim 7; Page 120-121; 130pp; English.
XX	
CC	The present sequence represents a p55 polypeptide. The specification
CC	describes p177, p88, p64, p55 and p46 polypeptides from <i>Neisseria</i>
CC	gonorrhoeae. The polypeptides are useful as vaccines, for preventing, or
CC	protecting a female patient against, <i>N. gonorrhoeae</i> colonization or
CC	infection. Such immunisation can prevent gonorrhoea in women. (Updated on
CC	29-AUG-2003 to standardise OS field)
XX	
SQ	Sequence 525 AA;

Query Match	97.2%;	Score 2643;	DB 5;	Length 525;
Best Local Similarity	97.3%;	Pred. No. 2.2e-233;		
Matches 511:	Conservative	5;	Mismatches 9;	Indels

RESULT 3
AAY75751
ID AAY75751 standard; protein: 507 AA.

XX
DT 12-SEP-2003 (revised)

XX DE Neisseria gonorrhoeae OPG 987 protein sequence SEQ ID NO: 2974

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX

KW antibacterial; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US009346.
 XX
 PR 01-MAY-1998; 98US-0083758P.
 PR 31-JUL-1998; 98US-0094869P.
 PR 02-SEP-1998; 98US-0098994P.
 PR 02-SEP-1998; 98US-0099062P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 25-FEB-1999; 99US-0121528P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ54513.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics.
 XX
 PS Claim 2; Page 1394; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to Neisserial
 CC bacteria (e.g. meningitis and septicemia), to detect the presence of
 CC Neisseria bacteria, or to raise antibodies. They may also be used to
 CC screen for agonists or antagonists, which may themselves have use as
 CC antibacterial agents. The polynucleotides of the invention may also be
 CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 507 AA;

Query Match 95.8%; Score 2606.5; DB 3; Length 507;
 Best Local Similarity 99.4%; Pred. No. 4.6e-230;
 Matches 505; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 18 MKTRSLISLLCLLSCSSWLPLEERTESRHFNTSKPVLDDNLIQIRHTPHNNGLSDIY 77
 Db 1 MKTRSLISLLCLLSCSSWLPLEERTESRHFNTSKPVLDDNLIQIRHTPHNNGLSDIY 60

Qy 78 LDDPHEAFAARAALIESAHSLSLQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 137
 Db 61 LDDPHEAFAARAALIESAHSLSLQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 120

Qy 138 NTRGLDLLALLDHPNIEVRLNPFVLRKWRALGYLTDFPRLNRRMHNKSFTADNRATI 197
 Db 121 NTRGLDLLALLDHPNI -VRLNPFVLRKWRALGYLTDFPRLNRRMHNKSFTADNRATI 179

Qy 198 LGGRNIGDEYFKVGEDTVFADLDLILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGNIG 257
 Db 180 LGGRNIGDEYFKVGEDTVFADLDLILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGNIG 239

Qy 258 KGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLISDSPAKGLDRDR 317
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 240 KGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLISDSPAKGLDRDR 299
 Qy 318 RKPIAGRLQDALKQPEKSVLYSPYFVPTKSGTDALAKLVQDGDVTVLTNSLQATDVA 377
 Db 300 RKPIAGRLQDALKQPEKSVLYSPYFVPTKSGTDALAKLVQDGDVTVLTNSLQATDVA 359
 Qy 378 AVHSGYVYKRPKLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFG 437
 Db 360 AVHSGYVYKRPKLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFG 419
 Qy 438 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVVTLDKHNRLQWHDPAIRK 497
 Db 420 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVVTLDKHNRLQWHDPAIRK 479
 Qy 498 TYPNEPEAKLWKRIAAKILSLPIEGILL 525
 Db 480 TYPNEPEAKLWKRIAAKILSLPIEGILL 507

RESULT 4
 AAY75753
 ID AAY75753 standard; protein; 508 AA.
 XX
 AC AAY75753;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 987 protein sequence SEQ ID NO:2978.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US009346.
 XX
 PR 01-MAY-1998; 98US-0083758P.
 PR 31-JUL-1998; 98US-0094869P.
 PR 02-SEP-1998; 98US-0098994P.
 PR 02-SEP-1998; 98US-0099062P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 25-FEB-1999; 99US-0121528P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ54515.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics.
 XX
 PS Claim 2; Page 1396-1397; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to Neisserial
 CC

```
CC bacteria (e.g. meningitis and septicemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols
XX
SQ Sequence 508 AA;

Query Match          95.1%; Score 2586; DB 3; Length 508;
Best Local Similarity 98.2%; Pred. No. 3.5e-228;
Matches 499; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 18 MKTRSLISLLCLLCCSSWLPPLLEERTESRHFNTPSKPVLDDNLIQIRHTPHNGLSDIY 77
Db 1 MKTRSLISLLCLLCCSSWLPPLLEERTESRHFNTPSKPVLDDNLIQIRHTPHNGLSDIY 60

Qy 78 LLDDPHEALAAARAALIESAHSLLDQYYIWRNDISGRLLFNLMYLAARGVVRLLDDN 137
Db 61 LLNDPHEAFAARAALIESAHSLLDQYYIWRNDISGRLLFNLMYLAARGVVRLLDDN 120

Qy 138 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMKNKFTADNRATI 197
Db 121 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMKNKFTADNRATI 180

Qy 198 LGGRNIGDEYFKVGEDTVFADLDIILATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 257
Db 181 LGGRNIGDEYFKVGEDTVFADLDIILATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 240

Qy 258 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 317
Db 241 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 300

Qy 318 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTLTNSLOATDVA 377
Db 301 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTLTNSLOATDVA 360

Qy 378 AVHSGYVKYRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 437
Db 361 AVHSGYVKYRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 420

Qy 438 SFNLDPRLNLTENGWVIESPKIAEQMERTLADTTPEYAYRVTLTKENRLQWHPDPAKTK 497
Db 421 SFNLDPRLNLTENGWVIESPKIAEQMERTLADTTPEYAYRVTLTKENRLQWHPDPAKTK 480

Qy 498 TYPNEPEAKLWKRIIAKILSLPIEGLL 525
Db 481 TYPNEPEAKLWKRIIAKILSLPIESLL 508

RESULT 5
ID AAY75752
AC AAY75752 standard; protein; 508 AA.
XX
XX AAY75752;
XX
XX 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 987 protein sequence SEQ ID NO:2976.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
XX WO9957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
XX
XX 31-JUL-1998; 98US-0094869P.
```

```
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Testein H, Venter JC;
XX
XX WPI; 2000-062150/05.
XX N-PSDB; AAZ54514.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics.
XX
XX Claim 2; Page 1395; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
XX represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ54673 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of the
XX invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the manufacture
XX of medicaments for treating or preventing infection due to Neisserial
XX bacteria (e.g. meningitis and septicemia), to detect the presence of
XX Neisseria bacteria, or to raise antibodies. They may also be used to
XX screen for agonists or antagonists, which may themselves have use as
XX antibacterial agents. The polynucleotides of the invention may also be
XX used in gene therapy protocols
XX
XX Sequence 508 AA;

Query Match          94.6%; Score 2573; DB 3; Length 508;
Best Local Similarity 97.8%; Pred. No. 5.5e-227;
Matches 497; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 18 MKTRSLISLLCLLCCSSWLPPLLEERTESRHFNTPSKPVLDDNLIQIRHTPHNGLSDIY 77
Db 1 MKTRSLISLLCLLCCSSWLPPLLEERTESRHFNTPSKPVLDDNLIQIRHTPHNGLSDIY 60

Qy 78 LLDDPHEALAAARAALIESAHSLLDQYYIWRNDISGRLLFNLMYLAARGVVRLLDDN 137
Db 61 LLNDPHEAFAARAALIESAHSLLDQYYIWRNDISGRLLFNLMYLAARGVVRLLDDN 120

Qy 138 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMKNKFTADNRATI 197
Db 121 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMKNKFTADNRATI 180

Qy 198 LGGRNIGDEYFKVGEDTVFADLDIILATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 257
Db 181 LGGRNIGDEYFKVGEDTVFADLDIILATGVSVEVSHDFDRYWASHAHNATRIIRSGDIG 240

Qy 258 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 317
Db 241 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 300

Qy 318 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTLTNSLOATDVA 377
Db 301 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTLTNSLOATDVA 360

Qy 378 AVHSGYVKYRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 437
Db 361 AVHSGYVKYRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 420

Qy 438 SFNLDPRLNLTENGWVIESPKIAEQMERTLADTTPEYAYRVTLTKENRLQWHPDPAKTK 497
Db 421 SFNLDPRLNLTENGWVIESPKIAEQMERTLADTTPEYAYRVTLTKENRLQWHPDPAKTK 480
```



```
QY 498 TYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TYPNEPEAKLWKRIAAKILSLPIEGLL 508

RESULT 6
ADL04719
ID ADL04719 standard; protein; 543 AA.
AC ADL04719;
XX
DT 06-MAY-2004 (first entry)
DE M. catarrhalis protein #485.
XX
KW Moraxella catarrhalis; infection.
XX
OS Moraxella catarrhalis.
XX
PN US6673910-B1.
XX
PD 06-JAN-2004.
XX
PF 04-APR-2000; 2000US-00540236.
XX
PR 08-APR-1999; 99US-0128416P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
XX WPI; 2004-178127/17.
DR N-PSDB; ADL02799.
XX
XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.
XX
XX Disclosure; SEQ ID NO 2405; 429pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents the amino acid
CC sequence of a M. catarrhalis protein.
XX
XX SQ Sequence 543 AA;

Query Match 30.6%; Score 831.5; DB 8; Length 543;
Best Local Similarity 35.9%; Pred. No. 6.3e-67;
Matches 194; Conservative 102; Mismatches 198; Indels 47; Gaps 12;

QY 22 SLISLLCLLLSCSSWLP-----PLEERTSRHPTNPKPVLNLLNIL-----QIRHT 67
Db 13 SIHTVMLVLLVGCKN-LPDTPLPKSLALTRQVQERHQMVKHGGSDGLISAINAQSHHH 71
QY 68 PHNNGLSDIYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLFNLMYLAABRG 127
Db 72 PNQSG---YPIITGANAFARSTLTDVANOSIDIQYIWHNDAGQLMLKDLWEAADRG 128
QY 128 VVRVRLLLDD-NNTRGDLLLALDSHPNIEVRLFPVLRKWRALGYLTDPPRLNRRMHN 186
Db 129 VIVRLLLDDFNSPELDQLLLRIKSKKNIAVRLINPMPYRGRFSRLNMLHPIRTNRMHN 188
QY 187 KSFTADNRATILGRNIGDEYFKVGEDTVFADLLILATGVSVGVSVDHFDRIYASHSAHN 246
Db 189 KSMTFDNKISVIGRNIYGNELNPNVHFNADLDVMLVGHVVGKITQTSFEIYWASPLSFD 248
QY 247 ATRIIRSGNIG--KGLQALGYND-----ETSRHALLRYEETVEQSPLYQKIQTGRIDW 297
Db 249 IETLVKHNKDDISGIKPIKVPFDELEKVKDSSDADRELRTYRQAMQNSTIGDQLLAQVVPF 308
```

```
298 QSVOTRLISDTPAKGLDRDRKPPPIAGRLQDALQKPEKSVVLVSPYFVPTKSGTDAKL 357
Db 309 FWTEDLIADNVIKLNGHSDNEFVSQLQOLQOPSKKUSIISYFVPTREGIDTLTL 368
QY 358 VQDGIDVTLTNSLQATDVAHVSGYVYKRPKLKAGIKLYELOPN-HAVPATKDK--GL 414
Db 369 AKMGVEVRILTNSPDATDVGIVHAGYAHWRKQLLAAGVHLFEIKSSAQSIQDNENRFPWT 428
QY 415 TGSSTVSLHAKTFIVDGKRIFIGSFNLDPPSRARLNTMGVIVESPKIAEQHERTIADT-- 472
Db 429 RQHSTTSLHAKAFVDDQIFIGSYNDPPRSANITELGVLIKDSKLAGQLHKLALSQA 488
QY 473 TPEYAVRVTLDKHNRLQWHDPATRKTPN-----EPEAKLWKRIAAKILSLPIEGL 524
Db 489 ITHQAYELKLDKAGNINWH-----TIENGOPVILHHEPHNMHVKDRIRIWLGLMPDWL 542
QY 525 L 525
Db 543 L 543

RESULT 7
ABO75163
ID ABO75163 standard; protein; 564 AA.
XX
AC ABO75163;
XX
DT 29-JUL-2004 (first entry)
DE
XX
XX Pseudomonas aeruginosa polypeptide #7338.
KW
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX
XX N-PSDB; ABD08734.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 23909; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
```

[illegible]

[illegible]

XX WPI; 2003-576092/54.
 DR N-PSDB; ADA30178.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 PS Example; SEQ ID NO 5591; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 SQ Sequence 275 AA;

Query Match 17.6%; Score 478.5; DB 6; Length 275;
 Best Local Similarity 41.0%; Pred. No. 5.9e-35;
 Matches 100; Conservative 51; Mismatches 78; Indels 15; Gaps 5;
 QY 294 RIDMQSVQTRLSIDTPAKGLDRDRKPPAGRLDALKQEKSVYLVSPYFVPTKSGTDA 353
 DB 35 KFDW--VAAEVVKDPSKIRSKAKKEHLNLFQNLHLEKPSNVLDLSAYFIPEKQAKI 92
 QY 354 LAKLVQGDIDVTLTNSIQATDVAHVHSGYVYKRPKLLKAGIKLYELQP-----NH 404
 DB 93 LSTLAKEGVEVRVLTNSFKANDVAVHAFYCKYKELLKNGVQLYELFPTPKEDLNKNT 152
 QY 405 AVPATKQK-GLTGSVSTSLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVWIESPKIAE 463
 DB 153 DELATKAKVNMKGLSRSSLTKLMLD-EQVFIGSFNFDPRSAYLNTGIVLDSPSLAK 211
 QY 464 QMERTLADTTPPEYAYRYVTLQKHNLQWHD--PATRKTYPNEPEAKLWKRTAAKILLSLPI 521
 DB 212 TIHTMDENLNKYAYKULQDPNNHIYWOQETPKGPVIYKKEPKMKWQKAGMKLLSWLPL 271
 QY 522 EGLL 525
 DB 272 EGFM 275

RESULT 13
 ID ABM88218 standard; protein; 2519 AA.
 XX
 AC ABM88218;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress responsive polypeptide SEQ ID NO:6464.
 XX
 KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
 XX
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314662P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX

PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer I, Provart N, Ricke D, Zhu T;
 XX
 XX WPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 PS Claim 1; SEQ ID NO 6464; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX
 SQ Sequence 2519 AA;

Query Match 15.4%; Score 419; DB 7; Length 2519;
 Best Local Similarity 28.2%; Pred. No. 5.5e-28;
 Matches 147; Conservative 76; Mismatches 186; Indels 112; Gaps 18;
 QY 32 CSCSSMLPPEER--TES-----RHFNTPKPVLLDNIQIRHPHNNGLSDIYLLDDPHE 84
 DB 377 CAQDSPLRALAGRAFTSTPEQPRHYAT-----LLDEGEG 411
 QY 85 ALAARAALTESAEHSLDLOYIWRNDISGRLLFNLMYLAERGVVRVLLDDNNTRGLDD 144
 DB 412 ALVARLNLRSATRSIDLQTYIFDKDSDARLVDELAAASRRGVKRVLDQLSAISDLQ 471
 QY 145 LLLALD-SHPNIEVRILFNP--FVLKRWALCY----LTDFFRLNRHMHKSFSTADNRATIL 198
 DB 472 ILGALSGAQNFQLRVYNPTFGKARLNYFDYAGSVLCCFRFRNORMHKNLLVDDAIGVV 531
 QY 199 GGRNIGDEYKVGEDTVFADLDILATGSSVVGVESHDFDRIYWASHASHAHNATRIIRSGNIGK 258
 DB 532 GGRNYQDDYYDMDREYNFRDRDLVIAGPEARAMAANFADFHWARRSVPAERL---NDVGR 588
 QY 259 GLQALGYNDETSRHALLRYRETYEQSPLYQKIQTGRIDWQS-----VQTRLSIDT 308
 DB 589 TLLREGV--FTLPPASFRFRVOR----VSAENDMDVFSRFSVDTALPVASVRYVADL 642
 QY 309 PAKGLDRDRRKPPFIAG-----RLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVQ--D 360
 DB 643 PRKH-RREKADAPLAGQHVTEPQLDALIAGAEVILQTPYLVLSPAQKLFRELKRPQ 701
 QY 361 GIDVTVLTNSLQATDVAHV-----SGYVYKRPKLLKAG 394
 DB 702 PPRVVVSSNSLAATDNPVYVLSYNSNRDRDEHPAPTQGLAVNPGTGSAYQNAERERRAA 761
 QY 395 IKL-----YELQP-----NHAVPATKDKLGTSSVTSLSHAKTFIVDGKRIFTGSPN 440
 DB 762 SEVETRLLTETRPSPFLGSAVKNKPLPVTTRKGRMG-----LHAKSLVDRRIIGVVGTHN 816
 QY 441 LDPRSARLNTMGVWIESPKIAEQM-ERTLADTTPPEYAYRV 480
 DB 817 FDRSRNNTYEGAVIIDDPAFAEQLAESILRDHPQNSWTV 857

RESULT 14
 ADA36398

[illegible]

RESULT 16	
ADG30702	
ID	ADG30702 standard; protein; 486 AA.
XX	
AC	ADG30702;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	Xanthomonas axonopodis pv citri plant pathology-related XAC0014 protein.
XX	
KW	Xanthomonas microorganism; plant; pathology; bacterial pest; Xac; Xcc;
KW	XAC.
XX	
OS	Xanthomonas axonopodis pv. citri.
XX	
PN	WO2003089647-A1.
XX	
PD	30-OCT-2003.
XX	
PF	22-APR-2003; 2003WO-BR000060.
XX	
PR	22-APR-2002; 2002US-0374620P.
XX	
PA	(AMPA-) FUNDACAO AMPARO A PESQUISA DO ESTADO.
XX	
PI	Da Silva ACR, Farah SC, Quaggio RB, Reinach FDC, Ferro JA;
PI	De Oliveira JCF, De Laia ML, Setubal JC, Furian LR;
XX	
DR	WPI; 2003-865444/80.
DR	N-PSDB; ADG30701.
XX	
PT	New nucleic acid molecule from a Xanthomonas microorganism, useful in
PT	determining the presence of Xanthomonas bacteria in a sample.
XX	
PS	Claim 8; SEQ ID NO 59; 145pp; English.

PS Claim 8; SEQ ID NO 59; 145pp; English.

The invention relates to a novel isolated nucleic acid molecule from a Xanthomonas microorganism where the nucleic acid molecule is associated with pathogenicity caused by the Xanthomonas microorganism, or its variant, that causes reduced or enhanced pathogenicity. The nucleic acid of the invention may be useful in detecting the presence of Xanthomonas bacteria in a sample, as well as in plant pathology, for example, for identifying nucleic acid molecules and proteins involved in pathology caused by bacterial pests. The current sequence is that of the Xanthomonas axonopodis pv. citri (Xac) plant pathology-related XAC protein of the invention.

Sequence 486 AA;

Query Match 10.9%; Score 296; DB 7; Length 486;
Best Local Similarity 25.2%; Pred. NO. 8.4e-18;
Matches 123; Conservative 63; Mismatches 209; Indels 94; Gaps 14;

QY 33 SCSSWLPLLEERTESRHFNTSKPVLLDNLIQLIRHTPHNGSLSDIYLDDDPHEALARAAL 92
||| ||| : || | | | : ||
Db 89 SYSSVCPDADCTE-----LAKIAQSTTGAPSSATEVHVLVDGAATYAALIEA 137

QY	93	IESAEHSLDQYIWRNDISGRLLFNLMYLAAEAGVVRVRLLLDDNNTRGLDLLLLALDOSH	152
Db	138	IRGARDHITHLEYIIFQPDHSGTAICAALMERARAGVKVRLMLDAIGSSAMTRRALRTLRE	197
QY	153	PNIEVRLFPFVLKWRALGYLTD- -PLNRRMHKSFTADNRATILGGRNIGDEYFKV	210
Db	198	AGVETAWFHPSQL- - - - -LKPFKRPWLNLRTHRKVIVIDGRIGFTGGINVTDDNEQ	249
QY	211	GEDTVFADLDILATGSSVVGVEYSHDFDRYWASHSAHNATRIIRSGNIGKGLQALGYNDETS	270
Db	250	VRKQAYRDLHVRQLGHVRSQLVFLDNL- - - - -YATSQGRAAPHGQOL- - - - -WPDDMP	300
QY	271	RHALLRYBETVEQSPLYOKIQTGRIDWQSVOTRLISDTPAKGLDRDRKPPITAGRLQDAL	330
Db	301	TRA- - - - -QGTVDQOVL- - - - -VSGPDSSWEAIIHRLMVA- - - - -AI	331
QY	331	KQPEKSVYLVGPFYFPTKSGTDAKLAKLVQDGDIVTLTNSLQATDVAAVHSGYVKYRKL	390
Db	332	HEAKHRVWLVTYPFVPGEAERWALTSAALGGLDVRLVLP- -RVSDSLVTVYAARSYFDEL	389
QY	391	LKAGIKLYELOPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIGSFNLDPPRSARLNT	450
Db	390	LEAGVRIYEYGR- - - - -MLHTKALLADDVDCVIGSANSFDSRSFRLNF	432
QY	451	EMGVVIESPKIAEQWERTLADTTPEYAVRVLTDKHNLQWHDPATRKYTPNEPEAKLWKR	510
Db	433	ELSMFLFRDQAVAAEM-AGLIGTDLQQAQEQVQFARHRPL- -WRS- - - - -RUPEA- - - - -	477
QY	511	IAAKILSLIL	519
Db	478	-FARLLSPL	485

RESULT 17

ABB49813
ID ABB49813 standard; protein; 504 AA.

AX
AC ABB49813;

XX
DT 05-FEB-2002 (first entry)XX
DE Listeria monocytogenes protein #2517.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX

AA
OS
Listeria monocytogenes.

XX PN WO200177335-A2.

XX
PD 18-OCT-2001

XX DE 11 2008 2001

[illegible]XX
11-AFK-2000; 2000FK-00004623.PA (INSP) INST PASTEUR.
XX

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fahi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cottart P;
PI Daniels J, Gobel W, Kref J, Kuhn M, Ng E, Vazquez-Bolland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX R202 M, Voss H;
XX WPI: 2002-010914/01.
DR

XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
PT and prevention of *Listeria* and related bacterial infections, and related
PT polypeptides.

Db 399 EAGVNVVHY-----DNFG-----LHSTLVIDDEVASVGTANMDNRSTLNF 441
Qy 452 MGVIKSPKIAEQMERTLAD-----TPPEYAYRVTLDK 484
Db 442 VNAFYDEGVARSLSQAFINDMKLSNKLTSSEYAKRNLLVK 482

RESULT 19

ADS04559
ID ADS04559 standard; protein; 495 AA.

AC ADS04559;

DT 04-NOV-2004 (first entry)

DE Staphylococcus epidermis polypeptide seqid 3854.

XX antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system.

OS Staphylococcus epidermidis.

XX US2004147734-A1.

XX 29-JUL-2004.

XX 01-DEC-2003; 2003US-00724972.

XX 08-NOV-1997; 97US-0064964P.

XX 13-AUG-1998; 98US-00134001.

XX 29-NOV-1999; 99US-00450969.

PA (DOUC/) DOUCETTE-STAMM L.

PA (BUSH/) BUSH D.

XX Doucette-Stamm L, Bush D;

XX WPI; 2004-580138/56.

DR N-PSDB; ADS00787.

XX New isolated polypeptide and encoding nucleic acid derived from

PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or

PT treating an S. epidermidis bacterial infection.

XX Claim 17; SEQ ID NO 3854; 741pp; English.

CC The invention describes an isolated nucleic acid comprising a nucleotide
CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC given in the specification. Also described are: a recombinant expression
CC vector; a cell comprising a recombinant expression vector of (1);
CC producing an S. epidermidis polypeptide; an isolated nucleic acid
CC comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection, comprising a nucleic acid cited above and a carrier; treating
CC a subject for S. epidermidis infection; a recombinant or substantially
CC pure preparation of an S. epidermidis polypeptide or its fragment; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection; detecting the presence of a Staphylococcus nucleic acid in a
CC sample; a computer readable medium having recorded in it the nucleotide
CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC system for identifying fragments of the Staphylococcus genome of
CC commercial importance; a computer based system for identifying fragments
CC of the Staphylococcus plasmids of commercial importance; identifying
CC commercially important nucleic acid fragments of the Staphylococcus
CC genome and/or plasmids; and identifying an expression modulating fragment
CC of the Staphylococcus genome and/or plasmids. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of an Staphylococcus epidermidis bacterial
CC infection. This is the amino acid sequence of a S. epidermis protein of

CC the invention.

XX Sequence 495 AA;

Qy Query Match 10.7%; Score 291; DB 8; Length 495;
Best Local Similarity 21.5%; Pred. No. 2.5e-17;
Matches 112; Conservative 79; Mismatches 166; Indels 164; Gaps 17;

Qy 1 MHTDPKIQAMPSETISP-----MKTRSLISLLCLLSCSSWLPPLPTESTRFNTSKP 55
Db 89 MIVDEQLEALKQDFSGKNHQIVKFKEMVQMLL-----YNNAAF 127

Qy 56 VLLDNILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAHSLSLDQYYIWRNDISGR 115

Db 128 LTTDN-----DLTIYDGHOKFDDLLINDIRHAQSYIHQYIIHSDNLGKQ 173

Qy 116 LFNLMYLAERGVRVRLLLDDNNTGRGLDLLLALDHPNIEVRLFNPFVIRKWRALGYLT 175

Db 174 LLHELEKKAEGIEVKWLYDDMGSRDL-----RKDKLKKFKQKGCHA 215

Qy 176 D--FP-----RLNRRMHNKSTADNRATILGGRNIGDEYFKVGEDTVFA---DL 222

Db 216 ESFFPSKLPILINLRMNNRNRKIVIDGTIGYVGGFNVGDEY--IGSKKFGYWRDTHLR 273

Qy 223 ATGSVVGEVSHDFRYWASHSAH-NATRIIR-----SGNIGKGLQALGYNDTSRHAL 274

Db 274 IKGDVNALQLRFLDWNQSOTRDLTYESYFPDVSDDGTIGIQAASSGPDE----- 326

Qy 275 LRYRETVESPLYQKIQGTGRIDWQSVQ---TRUISDTPAKGLDRDRRKPPIAGRLQALK 331

Db 327 -----DWEIKYGLKMISSA----- 342

Qy 332 QPEKSVLYSPYFVPTKSGTDALAKLVQDGIDVTLTNSLQATDVAHVHSGYVYKRP 391

Db 343 --KESYIOTSPYFIPQAFILDSIKIALGQVDVNMVFNKR--DHPFVYVATLKNVASLL 398

Qy 392 KAGIKLYELQPNHAPVATPKDGLTSGSVTSLHAKTFIVDGKRFIFGSFNLDPRSARL 451

Db 399 EAGVNVVHY-----DNFG-----LHSTLVIDDEVASVGTANMDNRSTLNF 441

Qy 452 MGVIKSPKIAEQMERTLAD-----TPPEYAYRVTLDK 484

Db 442 VNAFYDEGVARSLSQAFINDMKLSNKLTSSEYAKRNLLVK 482

RESULT 20

ABB49993

ID ABB49993 standard; protein; 482 AA.

XX ABB49993;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #2697.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

XX WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR001118.

PR 11-APR-2000; 2000FR-00004629.

PA (INSP) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Faihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Krefit J, Kuhn M, Ng E, Vazquez-Boland JA;

[illegible]

RESULT 25

ABB32493
ID ABB32493 standard; protein; 502 AA.

DT 23-JAN-2002 (first entry)

DE Staphylococcus aureus polypeptide SEQ ID NO 90.

Staphylococcus aureus; antimicrobial; growth rate; infection; vaccine; immunity.

OS Staphylococcus aureus.

PN WO200177365-A2.

PD 18-OCT-2001.

06-APR-2001; 2001WO-US011177.

PR 06-APR-2000; 2000US-0266327P.

PA (PHAA) PHARMACIA & UPJOHN CO.

PI Tomich CC;

DR WPI; 2002-010920/01.

DR N-PSDB; ABA51669.

PT Identifying antibacterial agents which bind to polypeptide encoded by
PT essential coding sequences of *Staphylococcus aureus*, and which reduce
PT growth rate of bacteria, by determining binding of agent to the
PT polypeptide.

PS Example 2; Fig 2: 67pp; English.

The invention relates to identifying antimicrobial agents that target the proteins (ABB32471-ABB32494) encoded by genes (ABA51623-ABA51634 and ABA51659-ABA51670) essential for the survival of *Staphylococcus aureus*, especially *Staphylococcus aureus*, including antimicrobial agents that interfere with the expression of essential coding sequence products and antimicrobial agents that interfere with the function of essential coding sequence products. The genes and proteins are useful for identifying agents that decreases the growth rate of a microbe preferably *S. aureus*. The agents identified by the above method are bactericidal and will be useful for treating a subject infected with *S. aureus* or at risk of being infected by *S. aureus* and may be used in vaccines for producing immunological response to *S. aureus* resulting in immunity. The present sequence is that of a protein encoded by an essential *S. aureus* polynucleotide coding sequence for expression in *E. coli*.

Sequence 502 AA;

Query Match

Best Local Similarity 21.1%; Pred. No. 6.6e-17;

Matches 106; Conservative 86; Mismatches 169

Qy	42	EERTSRHFTNSKPVLLD-----NILQRTHPHNNGLSYVLLDDPDHEALAAARALIES	95
Db	105	DNKQVKHHDLVRMLMDQDGFLTENNKNVDHFIDGNLDYDQVLKD-----IKN	152
Qy	96	AEHSJLDQYYIWRNDISGRLLFNLMYLAAERGVRVRLLLDNNTRGLDDLALLLDSHPNI	155
Db	153	AKEYIHUEYYTFALDGLGKRILHALEBKLGQLEVKKIYDDVGSK-----NV	199
Qy	156	EVRLFNPFLRKWRALG-----YLTDP-----RLNRMHKNSFTADNRATILGSGNTGD	205
Db	200	KVANFDHF-----KSJLGEVEAFFASKPLPLNFMRMNNHRKIIVIDGQLGVYGFGNIGD	254
Qy	206	EYFKVGEDTVPADJLILATGSWGEVSHDFORYWASHSAHNATRII-----RSGNIK	258
Db	255	EYLGJLKLGYWRDTHLRIQGDAVALQLFILDWNS-QAHRPOPEYDVKYFPKGNGPLG-	312
Qy	259	GLOALGYNETSRHALLRYRTVQSPLYQKIQTGRIDMSVQ-----TRLSDTPAKGLDR	315
Db	313	-----NSPIQAASGPASPDWHQIEYGYTTWMISA-----	341
Qy	316	DRRKPPITAGRLQDALKOPEKSVYLVSPVFPTKSGTDALAKLVODGI DVTLTNSLQATD	375
Db	342	-----KKSIVLYQSPXFI PDNSYINAIKAATKSGVDVHLMIPC--XPD	381
Qy	376	VAVHSGVKVKRPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFFIVDKGRIF	435
Db	382	HPLYWATFNASDLLSSGVKIY-----TYENG-----IHSKMCCLI DDEIVS	424
Qy	436	IGSNLDPRSARLLTNGMWIESPKIAEQMERTLATITPEYAYRVVTLDKHNRLOWHDAT	495
Db	425	VGTANMDFRSEPLNFEVNAFYVDENLAKDL-----RVAYEHDITKSKOL-----	469
Qy	496	RKTYNPEPEAKLWKRIAAILS	517
Db	470	KESYANRPLSVKPKESIAKLVS	491

RESULT 26

ADC21402

ID ADC21402 standard; protein; 502 AA.

```
AC ADC21402;
```

18-DEC-2003 (first entry)

DE Staphylococcus aureus prote

XX KW antibacterial: gene therapy: vaccine: diagnosis

gene function; bacterial infection; antibiotic-resistant bacterium.

XX OS Staphylococcus aureus.

YY
PN W02003029484-A2

XX
PD
10-APR-2003

XX DE 100C-DES-BC

XX
BB 28-SEP-2001 2001WA-110030370

XX
PA (PHAA) PHARMACY & MEDICAL CO.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]

DR N-PSDB; ADC21401.

XX
25

PT preventing staphylococcal infections, comprises contacting a polypeptide
PT and an agent to form a mixture and determining whether the agent binds
PT the polypeptide.

PS Example 2; SEQ ID NO 84; 174pp; English.

XX The invention relates to methods of identifying an agent that binds a

CC polypeptide from *Staphylococcus aureus* by contacting a polypeptide and an

CC agent to form a mixture and determining whether the agent binds the

CC polypeptide. The method is useful in identifying agents that bind gene

CC products critical for the survival of microbes, preferably *staphylococcus*

CC microbes, including agents that interfere with the function of such gene

CC products. The method is used to identify new agents useful for treating

CC antibiotic-resistant bacteria. This sequence represents a protein from

CC *Staphylococcus aureus* of the invention when the corresponding gene is

CC cloned into the expression vector pQE-60 or pQE-70. This encoding gene

CC sequence corresponds to the essential gene shown in ADC21335.

XX Sequence 502 AA;

Query Match 10.5%; Score 286.5; DB 7; Length 502;

Best Local Similarity 21.1%; Pred. No. 6.6e-17;

Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

QY 42 EERTSRHFTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95

Db 105 DNKQVQKHDLVRMLMDQDGLFTENNKVDHFIDGNDLYDQVLKD-----IKN 152

QY 96 AEHSLDLQYIWRNDISGRLLFNMLYAAERGVVRLLDDNNTRGDLDDLALLDLSHPNI 155

Db 153 AKEVIHLEYTYTFALDGLGKRILHALEEKLGKQLEVILYDDVGSK-----NV 199

QY 156 EVRLFNPFVLRKWRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205

Db 200 KMANFDHF-----KSLGGEVAFPAFKLPLNLFNMRNHRKIIVIDQGLGVGGFNIGD 254

QY 206 EYFKVGGEDTVFADLLDILATGSGVGEVSHDFDRYVASHSAHNATRII-----RSGNIGK 258

Db 255 EYLGKGLGYWRDTHLRIOGDAVDALQRLFDLWNS-QAHRPQFEDYVKYFPKNGPLG- 312

QY 259 GLQALGYNDTSRHALLRYRETVEOSPLYQKIQTGRIDWQSVQ---TRLISDTPAKGLDR 315

Db 313 -----NSPQIAAGSPADWHQIEGYTRKIMS----- 341

QY 316 DRKPPIAGRLQDALKQPEKSVLYSPVPTKSGTDALAKLVODIGDVTVLTSLOATD 375

Db 342 -----KKSIVLOSPIFIDNSVINAIIKAAGSDVDVHLMIPC--KPD 381

QY 376 VAAHSGYVKYRKPILLKAGIKLYELQPNHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIF 435

Db 382 HPLVYWTAFSNASDLLSSGVKIY-----TVENGFI-----IHSKMKLIDDEIVS 424

QY 436 IGSFNLDPRLNTEMGVVIETSPKIAEQMERTLATDTPYAYRVTLDKHRLQWHDPAT 495

Db 425 VGTANMDFRSELPNFEVNAFYVDENLAKDL-----RVAYEHDITKSKQL-----T 469

QY 496 RKTYPNEPEAKLWKRIAAKILS 517

Db 470 KESYANRPLSVKFESLAKILS 491

RESULT 27

ADC25088

ID ADC25088 standard; protein; 502 AA.

XX ADC25088;

AC ADC25088;

XX 18-DEC-2003 (first entry)

DT S. aureus polypeptide #9 encoded by the essential coding region.

DE Binding agent; growth rate; microbe; virulence; vaccine;

XX Bacterial infection; antibiotic-resistant bacteria; bacteraemia;

KW septic shock; metastatic infection; endocarditis; arthritis;

KW osteomyelitis; pneumonia; abscess; skin rash; food poisoning;

KW multisystem dysfunction; toxic shock syndrome; antimicrobial;

antibacterial; immunosuppressive; cytostatic; antiinflammatory.

Staphylococcus aureus.

US2003087321-A1.

08-MAY-2003.

28-SEP-2001; 2001US-00966521.

28-SEP-2001; 2001US-00966521.

(TOMI/) TOMICH C C.

(QUIN/) QUINN C L.

(ARVI/) ARVIDSON S.

(MOTT/) MOTT J E.

(HARR/) HARRIS D W.

Tomich CC, Quinn CL, Arvidson S, Mott JE, Harris DW;

WPI; 2003-606127/57.

N-PSDB; ADC25087.

Identification of agent that binds polypeptide, for treating bacterial

infections, particularly caused by antibiotic-resistant bacteria,

involves mixing polypeptide and agent, and determining whether agent

binds polypeptide.

Example 2; SEQ ID NO 84; 122pp; English.

The invention discloses a method for identifying an agent that binds a

polypeptide which comprises mixing a polypeptide and an agent to form a

mixture and determining whether the agent binds the polypeptide. Also

claimed is a method for identifying an agent that decreases the growth

rate of a microbe, making a *Staphylococcus aureus* with reduced virulence,

which comprises altering a coding sequence in an *S. aureus* to introduce a

mutation, a vaccine composition comprising the *S. aureus* organism with

reduced virulence. The methods are useful for identifying an agent that

binds a polypeptide, useful in treating (e.g. as a vaccine) bacterial

infections, particularly those caused by antibiotic-resistant bacteria,

including bacteraemia, septic shock and serious metastatic infections

including endocarditis, arthritis, osteomyelitis, pneumonia, abscesses in

virtually any organ, skin rashes, food poisoning or multisystem

dysfunction, i.e. toxic shock syndrome. The inventive method identifies

agents that are unrelated to existing antimicrobials and that target

different aspects of *staphylococcal* invasion of and replication in the

host. The sequence presented is one of the *S. aureus* polypeptides of the

invention encoded by the essential coding region.

Sequence 502 AA;

Query Match 10.5%; Score 286.5; DB 7; Length 502;

Best Local Similarity 21.1%; Pred. No. 6.6e-17;

Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

QY 42 EERTSRHFTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95

Db 105 DNKQVQKHDLVRMLMDQDGLFTENNKVDHFIDGNDLYDQVLKD-----IKN 152

QY 96 AEHSLDLQYIWRNDISGRLLFNMLYAAERGVVRLLDDNNTRGDLDDLALLDLSHPNI 155

Db 153 AKEVIHLEYTYTFALDGLGKRILHALEEKLGKQLEVILYDDVGSK-----NV 199

QY 156 EVRLFNPFVLRKWRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205

Db 200 KMANFDHF-----KSLGGEVAFPAFKLPLNLFNMRNHRKIIVIDQGLGVGGFNIGD 254

QY 206 EYFKVGGEDTVFADLLDILATGSGVGEVSHDFDRYVASHSAHNATRII-----RSGNIGK 258

Db 255 EYLGKGLGYWRDTHLRIOGDAVDALQRLFDLWNS-QAHRPQFEDYVKYFPKNGPLG- 312

QY 259 GLQALGYNDTSRHALLRYRETVEOSPLYQKIQTGRIDWQSVQ---TRLISDTPAKGLDR 315

DE M. xanthus protein sequence, seq id 16463.
KW Transgenic plant; DNA replication; gene regulation; gene expression.
OS Myxococcus xanthus.
XX US6833447-B1.
FN 21-DEC-2004.
XX 10-JUL-2001; 2001US-00902540.
XX 10-JUL-2000; 2000US-0217883P.
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX New substantially purified Myxococcus xanthus nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
XX identifying mutations in a gene of interest, and for constructing
XX mutations in a gene of interest.

XX Example 2; SEQ ID NO 16463; 25pp; English.
XX The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.
XX The nucleic acid is useful for determining gene expression, identifying
XX mutations in a gene of interest, and for constructing mutations in a gene
XX of interest. Sequences given in records for SEQ IDs 9692-16825 represent
XX a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from USPTO
XX Sequence 420 AA;

XX Query Match 10.1%; Score 275; DB 9; Length 420;
XX Best Local Similarity 23.5%; Pred. No. 5.7e-16;
XX Matches 114; Conservative 65; Mismatches 179; Indels 128; Gaps 15;

QY 55 PVLIDNLIQIRHTPHNGL---SDIYLLDDPHEALAAALIESAEHSLDLYVWENDI 111
DB 38 PGVSGGLTRYLPFRHGVQGNACQLLDGVEAYPAMLEAIRGARYVMEYVMSDA 97
QY 112 SGRLLFNLMYLAERGVRVRLLD-----DNNTRGLDLDLALDSDHPNIEVR 158
DB 98 VGELEFGQALAEAAERGVHVXVLYDAVGSWTSRRSFFAGLRAGVD-----IR 144
QY 159 LFNPPVLKRWALGYLTDFFLRMRMINKSFTADNRTILGGRNIGDEYFKVGEDTVFAD 218
DB 145 AFKPFSLS--RGLRHLL-----RRDRKILVGDGEVAFVGGVNSAHWAPAEAGAWRD 196
QY 219 LDILATGSVVGVEVSHDFDRYVWASHSAHNATRIIRSGNIGKQLGALGYNDTSRHALLYR 278
DB 197 DVLRIGPVAHEUERCSATW-----RMMPOGRF-----HRLTRL 232
QY 279 ETVEQSPLYQIKIQTGRIDWQSVOTRLISDTPAKG-----LDRDRKPPFIAGRLQDALQKP 333
DB 233 ERLRNP-----PRRGAVGLVLSRRS--IHRAYLHAIRRA 267
QY 334 EKSVLVSPFVPTKSTDLAKLVQGDIVTLVNTSLQATDVAAVHSGVYKVKPLPKA 393
DB 268 RRSVLVAAAYFIIDRRVMALREAAARGVEVHLLNA--RSDHPILFEFMAFAVERLLGA 325
QY 394 GIKLYELQPNHAPVATKDKLTGSSVTSVSLHAKTFIVDGKIFTCGSFNLDPRSRALNTEMG 453
DB 326 GVRIFEWQ-----RGV-----LHAKTAVDGVWVGWTIGSFNLERLSLAFNHEVN 368
QY 454 VVIESPKIAEQMBERTLADTTPEYAYRVTLDKHNRQLQWHDPAIRKTYVNEPEAKLWKRRAA 513

DB 369 AVFADPRLGQQLDESRGDCGD-CREYTLAEFR-----RPLWQKLE 410
QY 514 KILSLL 519
DB 411 RALSLL 416

RESULT 32
AAW55249
ID AAW55249 standard; protein; 206 AA.
XX AC AAW55249;
XX 26-JUN-1998 (first entry)
XX H. pylori ORF 05cpl191lorf35 protein.

XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
XX identification; binding compound; bacteria; life cycle; activator;
XX inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
XX bacterium.
XX Helicobacter pylori.
XX OS
XX WO9737044-A1.
XX 09-OCT-1997.
XX 27-MAR-1997; 97WO-US005223.
XX 29-MAR-1996; 96US-00625811.
XX 02-APR-1996; 96US-00758731.
XX 25-OCT-1996; 96US-00736905.
XX 28-OCT-1996; 96US-00738859.
XX 06-DEC-1996; 96US-00761318.
XX (ASTR) ASTRA AB.
XX Smith D, Alm RA;
XX WPI; 1997-503122/46.
XX N-PSDB; AAV24658.

XX Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) -
XX useful in vaccines to treat or prevent H. pylori infection and for
XX diagnosis of H. pylori infection.
XX Claim 14; Page 492; 1145pp; English.

XX This sequence is a Helicobacter pylori protein of unspecified function.
XX The protein may be used in a vaccine to prevent or treat H. pylori
XX infection or to identify H. pylori polypeptide binding compounds, useful
XX as potential H. pylori life cycle activators or inhibitors. The DNA and
XX probes derived from it may be used for the identification of H. pylori in
XX a sample, and the diagnosis of H. pylori infection. Nucleic acid
XX sequences complementary to the DNA act as antisense sequences, and can be
XX used to prevent the translation of H. pylori mRNA. Antibodies against the
XX protein can be used in immunoassays to evaluate the abundance and
XX distribution of H. pylori-specific antigens. The genomic sequence of H.
XX pylori (ATCC 55679) was determined from overlapping contigs generated by
XX mechanically shearing the bacterial DNA. The sequences were analysed for
XX ORF of at least 180 nucleotides, and the predicted coding regions defined
XX by computer evaluation. To identify likely H. pylori antigens for vaccine
XX development, the amino acid sequences predicted from various ORF were
XX analysed for significant homology to other known or exported membrane
XX proteins. Having identified and determined the sequences of interest,
XX particular regions can be isolated from H. pylori by PCR amplification
XX for recombinant polypeptide production, e.g. in E. coli hosts
XX Sequence 206 AA;

XX Query Match 10.1%; Score 274.5; DB 2; Length 206;
XX

PA (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX N-PSDB; ABD09136.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 24311; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 584 AA;
XX
XX Query Match 9.9%; Score 268.5; DB 7; Length 584;
XX Best Local Similarity 24.2%; Pred. No. 3.8e-15;
XX Matches 107; Conservative 59; Mismatches 167; Indels 109; Gaps 14;
XX
QY 44 RTSEPHNTSKPVL---LDNLIQIRHTPHNNGSLDIYLLDDPHEALAAALIESAEHSL 100
Db 192 RRADLHPAVSGPVPWRSDN---RPFALLNDGVRFF---PRMLLA-----IEAAQRSI 237
QY 101 DIQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDNNTGRLDLDLALDASHNIEVRLF 160
Db 238 ELEYLVEDGHCALFLVALLDARRRGVAVRCLFDGFCGLGSAWQIRLREAGGELRLY 297
QY 161 NPFVLRKRALG---YLTDFPRLNRRMHKSFADNRATILGGRNIGDEYFK-VGEDTVF 216
Db 298 NPL---RWKLTGNLY-----RDHRKLLIVDGRLVGVGGAGITDEFWEPVSDVSAW 345
QY 217 ADDILATGSGVGEVSHDFRYWASHSAHNA-----TRIIRSGNIGKGLQALGYND 267
Db 346 REVNVMDGPVADWAALFERQWLACLEKAWKPREGMTLRLPPQPGAARGLGRVAYAD 405
QY 268 ETSRHALLRYRETVESQSPLYQKIQTGRIDWQSVQTRLISTPAKGLDRDRKPPAGRLQ 327
Db 406 -----ARQHRDILQSLVRALNGSRR-----426
QY 328 DALKQPEKSVLYSPVPPTKSGTDALAKLVQGDIVTTLNLSQATDVAAVHSGVVKYR 387
Db 427 -----IWLATPFLPTWKVRRLRKAQGVREVRLLLAG-RLTDHAPFYAQORYY 476
QY 388 KPLLKAGIKLYELQPNHAVPATKDKIGTSSVTSLSHAKTFIVDGKRFITGSFNLDRPSAR 447
Db 477 PRLLRAGVRIHEYQPRF-----LHLKMMVMD-DWVSVCSCNCFDHNLR 518
QY 448 LNTMGVVIESPKIAEQMERTL 469
Db 519 FNLDANLEALDDPDTNEAASL 540
RESULT 35
AB068379
ID ABO68379 standard; protein; 579 AA.
XX

AC ABO68379;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #554.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX N-PSDB; ABD01950.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 17125; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 579 AA;
XX
XX Query Match 9.9%; Score 268; DB 7; Length 579;
XX Best Local Similarity 24.5%; Pred. No. 4.1e-15;
XX Matches 111; Conservative 60; Mismatches 186; Indels 96; Gaps 13;
XX
QY 37 WLPLEERTESRHNTSKPVLDDNLIQIRHTP---HNGLSDIYLLDDPHEALAAALIE 94
Db 186 WRPWEALAAQVSCYKG--LKALVTRTETPLANN---RVRLJNGEASFEMFKAIS 240
QY 95 SAHSLSLDQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDNNTGRLDLDLALDASHPN 154
Db 241 AARQVTLVQFFIVRDALGRLQQLLLERAANGVEVFFLYDAIGSHALPHRYVVERLRQGG 300
QY 155 IEVRLNPNPVLKRWALGYLTDFPRLNRRMHKSFADNRATILGGRNIGDEYFKVGEDT 214
Db 301 VQMHGFS-----TGSGMLNRF-QVNFNHRKVVVDGECGFGVGHNVGVEY--LGEXP 350
QY 215 VFA---DLDLATGSGVGEVSHDF--DRYWASHSAHNAFRIIRSGNIGKGLQALGYND 269
Db 351 PLAPWRDTHMELRGFAVACIQESFADWTWATHSL-----395
QY 270 SRHALLRYRETVESQSPLYQKIQTGRIDWQSVQTRLISTPAKGLDRDRKPPAGRLQDA 329
XX

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:13:30 ; Search time 46 Seconds
(without alignments)
943.581 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAAILLLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	831.5	30.6	543	2	US-09-540-236-2405
2	790	29.0	564	2	US-09-252-991A-23909
3	493.5	18.1	271	2	US-09-328-352-5559
4	478.5	17.6	275	2	US-09-328-352-5591
5	409	15.0	321	2	US-09-328-352-7685
6	291	10.7	495	2	US-09-134-001C-4273
7	286.5	10.5	493	2	US-09-828-523A-22
8	286.5	10.5	502	2	US-09-828-523A-90
9	275	10.1	420	2	US-09-902-540-16463
10	268.5	9.9	584	2	US-09-252-991A-24311
11	268	9.9	579	2	US-09-252-991A-17125
12	259.5	9.5	494	2	US-09-134-001C-4176
13	252	9.3	166	2	US-09-328-352-7628
14	252	9.3	422	2	US-09-602-787A-78
15	252	9.3	500	2	US-09-848-726-2
16	246	9.0	508	2	US-09-543-681A-4717
17	241	8.9	409	2	US-09-134-000C-3664
18	232	8.5	442	2	US-09-710-279-1000
19	231.5	8.5	386	2	US-09-107-532A-4088
20	228	8.4	800	2	US-09-252-991A-24635
21	220	8.1	501	2	US-09-489-039A-10241
22	219	8.1	491	2	US-09-583-110-4158
23	216	7.9	415	2	US-09-134-000C-3595
24	215.5	7.9	419	2	US-09-489-039A-12789
25	215	7.9	518	2	US-09-107-433-3177
26	208	7.6	378	2	US-09-902-540-11860
27	149	5.5	213	2	US-09-902-540-13301

28	141.5	5.2	478	2	US-09-438-185A-437	Sequence 437, Appl
29	141.5	5.2	933	2	US-09-141-206-2	Sequence 2, Appli
30	135.5	5.0	933	2	US-09-949-016-9934	Sequence 934, Ap
31	134	4.9	922	2	US-09-141-206-6	Sequence 6, Appli
32	133.5	4.9	458	2	US-09-489-039A-11632	Sequence 6, Appli
33	133.5	4.9	932	2	US-08-968-752B-6	Sequence 6, Appli
34	133.5	4.9	932	2	US-09-536-224-6	Sequence 6, Appli
35	133.5	4.9	933	2	US-09-107-149-17	Sequence 17, Appl
36	130.5	4.8	272	2	US-09-902-540-16500	Sequence 16500, A
37	130	4.8	459	2	US-09-543-681A-5540	Sequence 5540, Ap
38	123	4.5	155	2	US-09-902-540-14817	Sequence 14817, A
39	122	4.5	183	2	US-09-107-532A-5960	Sequence 5960, Ap
40	121.5	4.5	933	2	US-09-107-149-2	Sequence 2, Appli
41	119	4.4	117	2	US-09-328-352-7334	Sequence 7334, Ap
42	111.5	4.1	961	2	US-09-328-352-7358	Sequence 7358, Ap
43	110	4.0	1036	2	US-08-968-752B-4	Sequence 4, Appli
44	110	4.0	1036	2	US-09-536-224-4	Sequence 4, Appli
45	109	4.0	510	2	US-09-540-236-3660	Sequence 3660, Ap

ALIGNMENTS

RESULT 1

US-09-540-236-2405
; Sequence 2405, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATI
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540, 236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2405
; LENGTH: 543
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2405

Query Match 30.6%; Score 831.5; DB 2; Length 543;

Best Local Similarity 35.9%; Pred. No. 2.1e-74;

Matches 194; Conservative 102; Mismatches 198; Indels 47; Gaps 12;

Qy	22	SLISLCLLCSSWLP-----PLEERTSRHFNTPSKPVLNDNII-----QIRHT	67
Db	13	SIHTVMLVLLVGCKN-LPDTPLPKSLALTQVQERHQMVKGGSSDGLISAINAQSHHH	71
Qy	68	PHNGLSDIYLLDDPHEALAAARALTESAEHSJDLQYIWRNDISGRLLFNLMVLAARG	127
Db	72	PNOSG---YYPITTGANAFAARSTLTDVANQSIDQYIWHNDEAGQLMKDLWEAADRG	128
Qy	128	VRVRLDD-NNTRGLDLLLLDSDHNPTEVRLNPFVLRKWRALGYLTDFFPLNRRMHN	186
Db	129	VI VRLDDFNSSPELDQLLRISKHNTAVRLINPMYPGRFSNLMHPITRNRRMHN	188
Qy	187	KSTADNRATILGRRNIGDEYFKVGEDTVFADLDILATSGVGEVSHDPPRYWASHAHN	246
Db	189	KSMTFONKISVIGRRNIGNEYNLNPVNFADLDVMLGVHVGKIQTSFYIYASPLSFD	248
Qy	247	ATRIIRSGNIG--KGQALGYND-----ETSRHALLRYRETVEQSPLYQKIQITGRIDW	297
Db	249	IELVLKHNKDDISGKPVVDFDELEKVKOSSDADRELRTYRQAMQNSTICQDILLAQQVPF	308
Qy	298	QSVQTRLI SDTPAKGLDRDRRKPPPIAGRLQDALKQPEKSVYLVSVFYVPFKSGTDLAKL	357
Db	309	FWTEIDLADNVIKLNGHSDPNEFLVSQLOTQLGQPSKKLSIISFYVPFKTGIDTLITL	368
Qy	358	VQDGIDVTVLTSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPN-HAVPATKDK--GL	414
Db	369	AKMGVEVRLTNSFDATDVGIVHAGYAHWRKQLLAAGVHLFEIKSSAQSIQDNENRFRWT	428

; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 78
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-78

Query Match 9.3%; Score 252; DB 2; Length 422;
Best Local Similarity 23.4%; Pred. No. 2.6e-16;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;

QY 67 TPNHNGLSDIYLLDDPHEALAAALIESAEHSLDLOYIWRNDISGRLLFNLMYLAER 126
Db 38 TGGNNG-----FYSDYRESLKRTAAIDEAEYIYVEIYIMAWDSYTPFFAALERAHR 92
QY 127 GVRVRLLED-----DNNTRGLDLLLLALDSDHPNIEVRLFPVLRKWRAL 171
Db 93 GVKVRLLFDFHVGSKYKPGYHRLKKELNRMGFAYLM-----LPLQWR-- 135
QY 172 GYLTDPRLRNRMHKNSFTADNRATILGGRN-IGDEY-----FKVGEDTVFADLDILATG 225
Db 136 ---RRFRPDLNRHKKMLIIDGHTAFMGSONLIAPSYLQKKNIKLGR--WKDLMLVELTG 190
QY 226 SVVGEVSHDFDRYWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSP 285
Db 191 PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 234
QY 286 LYKIQOTGRIDWQSVQTRILISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVVLYSPYFV 345
Db 235 LVQLIPSG-----LHTKYMIADPDDTTGNEALGVLGSSNLDIRSFGNLYEISLMIK 273
QY 346 PTKSGTDALAKLVQDGDIVTLNSLQATDVAHVSGYKYPKLLKAGIKLYELOPNHA 405
Db 274 PDSLEAVTSACYRGVTVELFVS--EQADQFAIDHAQSSYYQALLLEAGVKIYQF-PKPD 330
QY 406 VPATKDKGLTGSVTSLSHAKTFIVD-----GKRIF--IGSFNLDPRSARLNTMGVVIES 458
Db 331 V-----LHTKYMIADPDDTTGNEALGVLGSSNLDIRSFGNLYEISLMIK 375
QY 459 PKTAEQWERTLADTTPEYAYRVTLDKHNRLOWHDPATKTY 499
Db 376 GNLIHEL-NALTDRTYRTVSFKLTLDKNQORSW-----RRRY 410

RESULT 15
US-09-848-726-2
; Sequence 2, Application US/09848726
; Patent No. 6562607
; GENERAL INFORMATION:

; APPLICANT: NAMPOOTHIRI, Madhavan
; TITLE OF INVENTION: No. 6562607el Nucleotide Sequences Coding for the cls Gene
; FILE REFERENCE: 032301 WD 1191
; CURRENT APPLICATION NUMBER: US/09/848,726
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-848-726-2

Query Match 9.3%; Score 252; DB 2; Length 500;
Best Local Similarity 23.4%; Pred. No. 3.4e-16;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;

QY 67 TPNHNGLSDIYLLDDPHEALAAALIESAEHSLDLOYIWRNDISGRLLFNLMYLAER 126
Db 116 TGGNNG-----FYSDYRESLKRTAAIDEAEYIYVEIYIMAWDSYTPFFAALERAHR 170
QY 127 GVRVRLLED-----DNNTRGLDLLLLALDSDHPNIEVRLFPVLRKWRAL 171
Db 171 GVKVRLLFDFHVGSKYKPGYHRLKKELNRMGFAYLM-----LPLQWR-- 213
QY 172 GYLTDPRLRNRMHKNSFTADNRATILGGRN-IGDEY-----FKVGEDTVFADLDILATG 225
Db 214 ---RRFRPDLNRHKKMLIIDGHTAFMGSONLIAPSYLQKKNIKLGR--WKDLMLVELTG 268
QY 226 SVVGEVSHDFDRYWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSP 285
Db 269 PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 312
QY 286 LYKIQOTGRIDWQSVQTRILISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVVLYSPYFV 345
Db 313 LVQLIPSG-----PGYTTEPNLR-----MFNSIVHHAERLILCSPYFI 351
QY 346 PTKSGTDALAKLVQDGDIVTLNSLQATDVAHVSGYKYPKLLKAGIKLYELOPNHA 405
Db 352 PDSLEAVTSACYRGVTVELFVS--EQADQFAIDHAQSSYYQALLLEAGVKIYQF-PKPD 408
QY 406 VPATKDKGLTGSVTSLSHAKTFIVD-----GKRIF--IGSFNLDPRSARLNTMGVVIES 458
Db 409 V-----LHTKYMIADPDDTTGNEALGVLGSSNLDIRSFGNLYEISLMIK 453
QY 459 PKTAEQWERTLADTTPEYAYRVTLDKHNRLOWHDPATKTY 499
Db 454 GNLIHEL-NALTDRTYRTVSFKLTLDKNQORSW-----RRRY 488

RESULT 16
US-09-543-681A-4717
; Sequence 4717, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4717
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4717

Query Match 9.0%; Score 246; DB 2; Length 508;
Best Local Similarity 23.3%; Pred. No. 1.4e-15;
Matches 112; Conservative 84; Mismatches 163; Indels 122; Gaps 23;

Db 350 LHGKVAVIDDDMTSGNSLNDPLSLNSLEANLVIRDVAFNGQLYQHLLRELARRHCRRIIS 409

Qy 482 LDKHNR--LQWHDPATRKTYNPEP 503
 :| :| :
Db 410 -RRHARGYWRAFA---DFPRLP 429

RESULT 21
US-09-489-039A-10241
; Sequence 10241, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10241
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10241

Query Match 8.1%; Score 220; DB 2; Length 501;
Best Local Similarity 23.0%; Pred. No. 5.6e-13;
Matches 105; Conservative 63; Mismatches 177; Indels 112; Gaps 19;

Qy 74 SDIYLLDDPHEALARAALIESAHSISLDQYIWRNDISGRLLFNLMYLAERGVVRVLL 133
 :| :| :| :| :| :
Db 134 NQLQLLTESDDVMQALRTDIOLARHNIEWYFIWQPGMADSVAESLMAAARRGVHICRLM 193
 :| :| :| :| :| :
Qy 134 LDD--NNTRGLDLLLLADSHPNIEV-----RLFNPFVLKRWALGYLTDLPRLNRR 183
 || :| :| :| :| :
Db 194 LDSAGSVAFFRSPWAAMMRNAGIEVVEALKVNLKRVF-----LR-----RMDLR 237
 :| :| :| :| :| :
Qy 184 MHNKSFTADNRATTLGGNIGD-EYFK-----VGEDTVFADLDILATGSVGEVSHDFRY 238
 :| :| :| :| :| :
Db 238 QHRKXWMDNYIATGSMNMVDPFRYQDQSGVGQ---WIDLMARMEGPVATSMGIVYSCD 294
 :| :| :| :| :| :
Qy 239 WASHSAHNATRIIRSGNIGKGLQALGYNDTSRIALLRYRETVEQSPLYQIKTGRIDWQ 298
 :| :| :| :| :| :
Db 295 W---EIEGKRILPP---PPDVNTMPED-ASGHTI-----HTIASGP----- 330
 :| :| :| :| :| :
Qy 299 SVQTRLISDTPAKGLDRDRKPPIAGRLQDALKOPEKSVYLVSPYFVFTKSGTDALAKLV 358
 :| :| :| :| :| :
Db 331 -----GFPEDL---IHQALLTAAYAAKEHLIMTTPYFVPSDDLHHAICTAA 373
 :| :| :| :| :| :
Qy 359 QDGIDVTVLNLSQATDVAAVHSGVYKVRKPLLKAGIKLYELOPNHAVPATKDKGLTGSS 418
 :| :| :| :| :| :
Db 374 QRGVDSVSIILP--RKNDSLLYGVASRAFFFTTELLAAGVKIYQF-----EGGL---- 417
 :| :| :| :| :| :
Qy 419 VTSLHAKTFIVDGKRIPIFGSNLPRSARLNTMGVWLESPKIAEQMERTLADTTPEYAY 478
 :| :| :| :| :| :
Db 418 ---LHTKSVLDGSELVGTVNLDMRSLWLNFEITLVIDDAGFSD-----LAAVDDIYIS 470
 :| :| :| :| :| :
Qy 479 RVTLDKHNRLQWHDPATRKTYNPEPEAKLWKRIAKI 515
 :| :| :| :| :| :
Db 471 RSRL-----LDARLWLRP---LWQRIARL 493

RESULT 22
US-09-583-110-4158
; Sequence 4158, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A

Query Match	7.9%;	Score 216;	DB 2;	Length 415;	
Best Local Similarity	22.4%;	Pred. No. 1e-12;			
Matches	87;	Conservative 62;	Mismatches 154;	Indels 86;	Gaps 13;

Qy	47	SRHENT--SKPVLLDNIIQIRHTPHNNGSLDIYLLDDPHEALAAARAALIESAEHSLDLOY	104
Db	105	NRYFRNDASPLAKNSLQ-----LFTDGOEKFOALFEDIRAAKETVHVEY	150
Qy	105	YIWRNDISGRLLFNFLMAYLAABRGVRRLDDNNTRGLDDLLA--LDSPNIEVRLNFP	162
Db	151	YAFENDTIGNQFLDLVLEKLEHGEVEYLLYDPWGSFGANKKFFARYVDAGGVAFITSR	210
Qy	163	FVLKRWALGYLTDFPRLNRRMHNKSFTADNRATILGGRNIGDEYFKVGED--TVFADLDI	221
Db	211	DMIRK-----TRLNYHLHRKIIVVIDGKIGWTGGFNVDQYLVNTEKFGYWRDTHI	260
Qy	222	LATGSVVEGVSHDFRYWASHSAHNATRIIRSGNTGKGLQALGYNDETSRHALLRYETV	281
Db	261	RLVGTAVFSLOEIFIMDW-----NA-----SVKYPEER-----MTYHEKY	295
Qy	282	EOSPLYOKIQTRIDWQSVQTRLISDTPAKGLDRDRKPPAGRLQDALKQPEKSVYLVS	341
Db	296	FKLPEDHEV-----EHLSQLQVSDGP-----DSEBEILKSGFVRMIFSAEKSVMWQT	342
Qy	342	PYFVPTKSGTDALAKLVODGDVTVLTNSLQATDVAAVHSGYVKYKPKLLKAGIKLYELQ	401
Db	343	PYLLPDSMINALLVAVRSGVDVRIIMPCM--PDHPFIYRATQYYANYLHKGKIKIY---	397
Qy	402	PNHAVPATKDKGLTGSSVTSLHAKTFIVD	430
Db	398	-----IYDSGF-----IHAKTIWVID	412

RESULT 24
US-09-489-039A-12789
; Sequence 12789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12789
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12789

Query Match	7.9%;	Score 215.5;	DB 2;	Length 419;
Best Local Similarity	21.4%;	Pred. No. 1.2e-12;		
Matches	100;	Conservative 59;	Mismatches 204;	Indels 105; Gaps 11;

QY	76	IYLLDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNFLMYLAAEAGVRLILD	135
		: : : : : : : : : : : : : : :	
Db	16	ITLLENGDQYYPALFAAIGASRRVILIESFIWFEDVGRRLHAVLLFAARRGIQVEVLID	75
		: : : : : : : : : : : : : : :	
QY	136	DNNTRGLDLLLLALDSHPNIEVLFNFVLRKWRALGYLTDFPLNRMRMHMKSFTADNRA	195
		: : : : : : : : : : : : : : :	
Db	76	GYGSPDLSDFVGELTAAGVIFRYIDP----RPKLMGRNT---LFRMRHKIVVIDDTT	128
		: : : : : : : : : : : : : : :	
QY	196	TILGGRNIGDEYFKVGBDTPFADLIDILATGSVVGVEVSHDFRYWASHSAHNATRIIRSGN	255
		: : : : : : : : : : : : : : :	
Db	129	AFVGGINYSAEHMSDYGEAKQDVAVKVEGPVLIDL-QFE-----	168
		: : : : : : : : : : : : : : :	
QY	256	IGKGLQALGYNDSTSHALLRYRETVEQSPLYQIQIGRIDWQSVQTRLISDTPAK----	312
		: : : : : : : : : : : : : : :	
Db	169	----LENLP-NSETARKWRRRRHQPE-----INOTGEAAQL	201
		: : : : : : : : : : : : : : :	

```

313 QY 313 ---LDRDRKPPPIAQRQLQDALKEKSYYLVSPVPTKSGTDALAKLVQDGIDVTLTN 369
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
202 QY 202 FWNEDNQDHRDDIERHVLKMLTSARREVLIANAFFPGYRLLHANNAARGVRVKLIQ 261
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
370 QY 370 SLQATDVAAHVSHGVYKRKPRLKAGIKUYELOPNHAVPATXDKGLTGSSVTSIHKFTFV 429
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
262 QY 262 G--EPDIPVKFGAHLLHYHLVGKGVOIYEYRR-----PLHGKVALA 302
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 QY 430 DGKRIFITGSFNLDPRSLRNTEMGMWSTESPKIAEQMERTLATDTPEYAVRVTLDKHNRLQ 489
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
303 QY 303 DDHWATVGGSNLDPISLSLNLEANLIIHDRVFNOFLRDNLNGLIARDCORIDKTMLPKRN 362
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 QY 490 W-----HDPA-----TRKTYNEPEAKLWKRI 511
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
363 QY 363 WRRLGVSVMAFHLRHFPFWGVLPAPHTPRLARVSPVPQPELETQDRV 410
Db Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 25
US-09-107-433-3177
; Sequence 3177, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/107,433
  FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 60/ 085131
  FILING DATE: May 12, 1998
  APPLICATION NUMBER: 60/051553
  FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
  NAME: Ariniello, Pamela Deneke
  REGISTRATION NUMBER: 40, 489
  REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (781)893-5007
  TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3177:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 518 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHETICAL: YES
  ORIGINAL SOURCE:
    ORGANISM: Streptococcus pneumoniae
  FEATURE:
    NAME/KEY: misc feature
    LOCATION: (B) LOCATION 1...518
    SEQUENCE DESCRIPTION: SEQ ID NO: 3177:
US-09-107-433-3177

```

Query Match 7.9%; Score 215; DB 2; Length 518;
Best Local Similarity 21.2%; Pred. No. 1.9e-12;
Matches 100; Conservative 61; Mismatches 140; Indels 170; Gaps 20


```

; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9934

Query Match      5.0%; Score 135.5; DB 2; Length 933;
Best Local Similarity 18.9%; Pred. No. 0.0005;
Matches 123; Conservative 92; Mismatches 179; Indels 257; Gaps 34;

QY 30 LLCSSS-----WLPPEERTES-----RHFNSTSKP-----VL 57
Db 289 LILKSSYRQARWQAQETELAQGPGDRFLQHRHDSYAPPRPGTLARFWVNGAGYFAAV 348
QY 58 LDNILOIRHTPHNGLSDIYLLD---DPHEAALAAALIESAEHSLDLQYIWRNDISGR 114
Db 349 ADAILRAQ-----BEIFITDWLSP-EVYLKRP-----HSD--WRLDI--- 385
QY 115 LFNLMYLAERGVVRLLDDNNTRGLDGLLALDS-----HPNIEVRLNFPV 164
Db 386 ---MLKRAEGGVRSILLFKE-----VELALGINSYKRALMLLHPNIKV-MRHDPQ 435
QY 165 LRKWRALGYLTFPRLNRRMHNKSTADNRATILGG-----RNIGD----- 205
Db 436 VTLW-----AHHEKLLVVDQVVAFLGGLDLAYGRWDDLHYRLTDLGDSSES 481
QY 206 -----EYFKVGEDTFADLDILATGSSVGEVSHDF---DRYWAS 241
Db 482 AASQPTPRPDSPATPDLSHNQFFWLGKD---YSNL-----ITKDWQLDRPPEF 528
QY 242 HSAHNATRIIRSGNIGKGLQALGVNDETSRHALLRYRETVE-----QSLYQKIQTGRID 296
Db 529 FIDRETTPRMPWRDGVVHGLPARD-LARHFIQRWNETKTKAKVKTPIYPYLLPKS-- 585
QY 297 WQSVOTRLISDTPA-----KGLDRDRKPPPIAGRLQDA-----LKQPKSYL 339
Db 586 -TSTANQLPFTLPGGQCTTVQVLRSDVRWS-----AGTLENSILNAYLHTIRESQHFLYI 639
QY 340 VSPYFVPTKSGTDALAKLVQDGID-----VTVLT-----NSL 371
Db 640 ENQFISCDGRTVLNKVGEIVDRILKAHKQGCWYRVYVLLPLPGPEGDISTGGNSI 699
QY 372 QATDVAHVHSGY-----VKYRKLKAGIKLYELQPNHAVPATKDK 412
Db 700 Q---AILHFTYRTLRCRGEYSILHRLKAAMGTAWRDYISICGLRTHGELGHPV----- 749
QY 413 GLTGSSVTLHAKTFIVDGKRFIFGSEFNLDPRS--ARLNTMGVWIE---SPKI---AE 463
Db 750 ----SELIYIHSKVLIADDRVTIIGSANINDRSLGKRDSSELAVLIEDTETEPSLMNGAE 805
QY 464 QMERTLADTTPYAYRVTLDKHNR--LOWHDPATKTYPNEPEAKLWKRIA 512
Db 806 YQAGRFAISLRKHCFCGVILGANTRPDLDRDPICDDFF-----QLWQDMA 850

RESULT 31
US-09-141-206-6
; Sequence 6, Application US/09141206
; Patent No. 6187559
; GENERAL INFORMATION:
; APPLICANT: Steed, Paul M.
; APPLICANT: Lasala, Daniel J.
; TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A
; Patent No. 6187559
; FILE REFERENCE: 4-30148/PI/CGC1954/R
; CURRENT APPLICATION NUMBER: US/09/141,206
; CURRENT FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/057,802
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Human
```

```

US-09-141-206-6

Query Match      4.9%; Score 134; DB 2; Length 922;
Best Local Similarity 18.4%; Pred. No. 0.0007;
Matches 121; Conservative 84; Mismatches 173; Indels 278; Gaps 32;

QY 30 LLCSSS-----WLPPEERTES-----RHFNSTSKP-----VL 57
Db 289 LILKSSYRQARWQAQETELAQGPGDRFLQHRHDSYAPPRPGTLARFWVNGAGYFAAV 348
QY 58 LDNILOIRHTPHNGLSDIYLLD---DPHEAALAAALIESAEHSLDLQYIWRNDISGR 114
Db 349 ADAILRAQ-----BEIFITDWLSP-EVYLKRP-----HSD--WRLDI--- 385
QY 115 LFNLMYLAERGVVRLLDDNNTRGLDGLLALD-----SHNIEVRLNFP 162
Db 386 ---MFKRAEGGVRSILL-----FKELEALGINSYKRALMLLHPNIKV-MRHHP 433
QY 163 FVLKWRALGYLTFPRLNRRMHNKSTADNRATILGG-----RNIGD----- 205
Db 434 DQVTLW-----AHHEKLLVVDQVVAFLGGLDLAYGRWDDLHYRLTDLGDS 479
QY 206 -----EYFKVGED-----TV 215
Db 480 ESAASQPTPRPDSPATPDLSHNQFFWLGKDYNSLITKDWQLDRPPEFIDRETTPRMP 539
QY 216 FADLDILATGSSVGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGVNDETSRHALL 275
Db 540 WRDGVVHGLPARDLARHFIQRW-----NFTKTKA-----KYKTPTYPYLLP 583
QY 276 RYRETVESPLYQKIQTGRIDWQSVQTRLISDTPAKGLDRDRKPPPIAGRLQDA----- 329
Db 584 KSTANQLP--TLPGGQC--TTVQV-----LRSDVRWS-----AGTLENSILNAYL 627
QY 330 --LKQPKSYLVGPYFVPTKSGTDALAKLVQDGID-----VTVLT----- 368
Db 628 HTIRESQHFLYENQFFISCDGRTVLNKVGEIVDRILKAHKQGCWYRVYVLLPLPGF 687
QY 369 -----NSLOATDVAHVHSGY-----VKYRKLKAGIKLYEL 400
Db 688 EGDISTGGNSIQ-----AILHFTYRTLRCRGEYSILHRLKAAMGTAWRDYISICGLRTHGE 743
QY 401 QPNHAVPATKDKGLTGSVTSVLSHAKTFIVDGKRFIFGSEFNLDPRS--ARLNTMGVWIES 458
Db 744 LGHPV-----SELIYIHSKVLIADDRVTIIGSANINDRSLGKRDSSELAVLIED 793
QY 459 PKIAQOMERTLADTTPYAYRVTLDKHNR--LOWHDPATKTYPNEPEAKLWKRIA 512
Db 794 ---TETEPSLMNGAEYQAGSVILGANTRPDLDRDPICDDFF-----QLWQDMA 839

RESULT 32
US-09-489-039A-11632
; Sequence 11632, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11632
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11632

Query Match      4.9%; Score 133.5; DB 2; Length 458;
Best Local Similarity 20.0%; Pred. No. 0.00024;
```


MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,224
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/968,752
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Giotta, Gregory J.
REGISTRATION NUMBER: 32,028
REFERENCE/DOCKET NUMBER: ONYX2004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-222-9700
TELEFAX: 510-222-9758
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-536-224-6

Query Match 4.9%; Score 133.5; DB 2; Length 932;
Best Local Similarity 18.6%; Pred. No. 0.0008;
Matches 101; Conservative 80; Mismatches 165; Indels 197; Gaps 25

Qy	98	HSLDQYIWRNDISGRLLFNWLYAAERGVVRPFLILLDDNNTRGLDDLIALDS-----	151
Dd	377	: : : : : HSSD-----WLRI-----MLKKAEBEGVRSILFKB-----VELALGINSYSKRT	419
Qy	152	----HPNIEVLRFNEPVLRKWRLGYLTDPPRLNRMRHNKSFTADNRATITLGSRNIG--	204
Dd	420	: : : : : LMMLHPNIKV-WRHPDLVTIW-----AHHEKLIVVDQVVAFLGGLDLAFCR	464
Qy	205	--DEYFK--VGEDTVFADLDILATGS---VVGEVSHD-----FDR	237
Dd	465	: : : : : WDDVQVRYLTDLGDPSPEPVHLQTPTLGSDDPAATPDLSHQFFFWLKGDOYNLIITKDWVQLDR	524
Qy	238	YWASHASANNATRIIRSGNIGIKGLOALGVNDETSRHALLRYRETVE-----QSPLYOKIQOT	292
Dd	525	: : : : : PFEDIREDTTPRMWRDVGVVHVGAARD-LARHFTQRWNFTKTTKARYKTFELPYL---	581
Qy	293	GRIDQSVQTRLISDTP-----AKGLDRDRRKPPTAGRLOA-----LK	331
Dd	582	: : : : : LPKSTSTANMLPMIPGGQCATTQVULRSVDRWS-----AGLENSILNAYLHTIR	631
Qy	332	OPEKSYLVSPYPVPFKTSGTDALAKLVODGID-----VTULT-----	368
Dd	632	: : : : : ESQHFIYTIENQFPFISCSGDRTVLNKVGBEIVDRILKAHQEQCFRVVLLPLLPFGEGDI	691
Qy	369	-----NSLOATDVAAVHSYG-----VKYRPPLLKAGIKUYELOPNH	404
Dd	692	: : : : : STGGNSIQ-----AILHFTYRTLRCRBHSILHRUKAAMGTAWRDYMSICOLRTHBSLGGH	747
Qy	405	AVPATKDGLTGSSVTSLSHAKTFTVDGKRFIGFSNLDPBS--ARLNTSMGVVIESPKTA	462
Dd	748	: : : : : PI-----SELIYTHSKMLIADDTVTIIGSANINDRSLGKKRDSBLAIILKO----	793
Qy	463	EQUERTLADTTPEYAIRVTLDKHNR-----LOWHDPATRKTYTPNEPEAKLMK	509
Dd	794	: : : : : TEWESPMDMGVEYQAGRALSIRGCRFSVILGANTWPDLDRDPVCDDFF-----QLWQ	847
Qy	510	RIA	512
Dd	848	ETA	850

RESULT 35
US-09-107-149-17

```

; Sequence 17, Application US/09107149
; Patent No. 6274363
; GENERAL INFORMATION
;
; APPLICANT: Leung, David W.
; APPLICANT: Tompkins, Christopher K.
; TITLE OF INVENTION: PHOSPHATIDYLCHOLINE PHOSPHOLIPASE D
; FILE REFERENCE: 077319/0144
; CURRENT APPLICATION NUMBER: US/09/107,149
; CURRENT FILING DATE: 1998-06-30
; EARLIER APPLICATION NUMBER: 08/768,147
; EARLIER FILING DATE: 1996-12-17
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-107-149-17

```

Query Match	4.9%; Score 133.5; DB 2; Length 933;
Best Local Similarity	18.3%; Pred. No. 0.0008;
Matches	121; Conservative 88; Mismatches 175; Indels 277; Gaps 33;
QY	30 LLCSGSS-----WLPPLERTES-----RHFNTPSKP-----VL 57
Db	289 LIUKCSSYRQARWQAQBITELAQCPGRDFLQLHRHDSYAPRPGOTLARWFWNGAGYFAAV 348
QY	58 LDNILOTRHPTNNGSLDIYLLD---DPHEALARAALIESAEHSLDLQYYIWNRDISGR 114
Db	349 ADAILRAQ-----BEIFITDWLSP-EVYLKRP-----HSDD-----WRLDI--- 385
QY	115 LFLNLMYLAERGVVRVLLDDNNTRTGLDLLLALDS-----HPNIEVLRFNPFV 166
Db	386 ---MLKRKAEEGVRSILLFKE-----VELALGINSYKRALMLLHPNIKV-MRHDPD 435
QY	165 LRKWRALGYLTDPRLNRMRHNSFTADNRATILGG-----RNIGD----- 205
Db	436 VTLW-----AAHEKLLVVQVAVFLGGDLAYGRWDDLHYRLTDLGDSSES 481
QY	206 -----EYFKVGED-----TVFA 217
Db	482 AASQPTTPRDPSPATPDLSHNQFFWLKDYSNLITKDWQLDRPFDFIDRETPRMPWR 541
QY	218 DLBILATGSSVVGVEVSHDFRYWASHSAHNATRIIRSGNIGKGLQALGVNDETSRHALRY 277
Db	542 DVGVVVHGLPARDLARFIQRW-----NFTKTKA-----KYKTPITYLLPKS 585
QY	278 RETVEGSPKQKIQTRIDWQSOTVRLISDTPAKGLDRDRKPPIAGRLQDA----- 329
Db	586 TSTANQLPF--TLPGGQC--TTVOV-----LRSDVRWS-----AGTLENSILNAYLHT 629
QY	330 LKOPEKSVYLVSPYFVPTKSGTDALAKLVODGID-----VTVL----- 368
Db	630 IRGSQHPLYIENQFFISCSGRTVLNVKGVEIDVRLKAHKQGCWYRVYVLLPLLPGPEG 689
QY	369 -----NSIQATDVAHVSGY-----VKYRKPLLKAGIKLYELQP 402
Db	690 DISTGGGNSIQ-----AYLHFTYRTLRCGEYSILHRLKAAMGTAWRDFVISICGLRTHGELG 745
QY	403 NHAVPATKGLTGTSSVTSIHAKTFFIVDGKRIFGSFLNDPRS--ARLNTSMGVVIE--- 457
Db	746 GHPV-----SELIYIHSKVLIAADDRVTIIGSANINDRSLLGKRDSELAUVIETTE 795
QY	458 -SPKI---ABQMERTLATDTPPEYAYRVTLDKHNR--LQWHPDPAKTKTYPNPEAKLWKRI 511
Db	796 TERSLMNGAEYQAGRFAISLRKHCFCGVILGANTRPDLDRPDI CDDPF-----QLWQDM 849
QY	512 A 512
Db	850 A 850

RESULT 36

US-09-540-16500
; Sequence 16500, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16500
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-540-16500

Query Match 4.8%; Score 130.5; DB 2; Length 272;
Best Local Similarity 28.0%; Pred. No. 0.0002;
Matches 49; Conservative 22; Mismatches 73; Indels 31; Gaps 4;

Qy 299 SVOTRLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLV 358
Db 10 TVLARGIADGDEDFETLRTV-----LLGLATARTSVRIPTYFLPDTALITALSVA 63

Qy 359 QDGIIDVTVL-----TNSLQATDVAHVHSGYVYKRPKLLKAGIKLYELOPNHVPATKDKGLT 415
Db 64 LRGVQVDILLPEKGNLPLVQWAAWALW-----QVLRPGCRIFLTAPPFD----- 108

Qy 416 GSSVTSLSHAKTFIVDGKRIFTGSFNLDPKRSARLNTMGVIVIESPKIAEQMERTLA 470
Db 109 -----HTKLMVVDGWSLIGSANWDPKRSRLNFEEVCEYDTALAMQEGVVA 156

RESULT 37
US-09-543-681A-5540
; Sequence 5540, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR FILING DATE: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5540
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5540

Query Match 4.8%; Score 130; DB 2; Length 459;
Best Local Similarity 20.7%; Pred. No. 0.00054;
Matches 98; Conservative 70; Mismatches 196; Indels 110; Gaps 21;

Qy 47 SRHFTSKPVLNDILQIRTPH-----NNGLSDIYLLDDPHEALAAARAL---IESA 96
Db 1 TREIST---VWLSKLKAKHQHQAQLPKLAQSVADVETL---YQTSAPRSTLLKYIAQA 54

Qy 97 EHSLLDQYIWRNDSIGRLNLFNMLYA-----ABRGVVRVRLILDDNNT 139
Db 55 QKEIFITALLYLHDEAGEEILEALYTAQKQRPESLITVVDWHRAQRG-RIGVSADATNA 113

Qy 140 RGLDDLLALDSDHPNIEVRLFN-PFVLRKWRALGYLTDFPRLNRRMHNSFTADNRATIL 198
Db 114 ---DWYHKVAQHPDVPVIGYIPVNTRE--ALGVL-----HLKGFIFDD-TVIY 157

Qy 199 GGRNIGDEYFKVGBEDTVFADLDILATGVSIVGEVSHDFDRYWAHSAHNATR-----IIRSG 254
Db 158 SGASINNVYLH-----KLD-----KYRYDRYHIHNAELANTMKQFIVDSL 198

Qy 255 NIGKLGALGVNDSTSRHALLRYRETVEQSPLYQKIQTGRIDMOSVQTRILISDTPAKGLD 314
Db 199 LQSDAIQRL--DSEERVCAEIKNCIRQRFNRLRTHDGYDIKGNASNNQLTVPVLGLG 255

Qy 315 RDRKPPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDIDVTVLTNSLOAT 374
Db 256 ---KCNILNKTISHLMAVTEKLVICTPYFNLPAILVRQISHLRNGKQVEIIGDKTAN 312

Qy 375 DVAAVHSGYVYKRP--LLKAGIKLYEL-----QPNHAVPAPATKDKGLTGSS 418
Db 313 DF-----YIPPEPFKIIGALPYLYEINLRKTFQRFQFIDNDQLTVRLWKGDNT--- 363

Qy 419 VTSLSHAKTFIVDGKRIFTGSFNLDPKRSARLNTMGVIVIESP--KIAEQMERTLA 470
Db 364 ---YHLKGMVVDDEWQLITGNLNPRAWGLDLENAILLHDPHHELHEQRHKELA 414

RESULT 38
US-09-902-540-14817
; Sequence 14817, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14817
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14817

Query Match 4.5%; Score 123; DB 2; Length 155;
Best Local Similarity 30.2%; Pred. No. 0.00044;
Matches 38; Conservative 17; Mismatches 63; Indels 8; Gaps 2;

Qy 78 LDDDPHEALAAARAAALIESAHSLSLDQYIWRNDSIGRLNLFNMLYAAERGVVRLLDDN 137
Db 38 VLESRRDAYPAMLEAIDAARTTILSLCSYIFDNDMAGRRFVEALSAAVKGVEVRLVDAV 97

Qy 138 NTRGLDDLLALDSDHPNIEVRLFNPFVLRKWRALGYLTDFPRLNRRMHNSFTADNRATI 197
Db 98 GSRYTWPPIILGRLLRRAGVRAARFLP-SLMPYR-----LPFANLRNHRKLMVVDGRVGF 149

Qy 198 LCGRNI 203
Db 150 TGGMNI 155

RESULT 39
US-09-107-532A-5960
; Sequence 5960, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynnn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:24:35 ; Search time 166 Seconds
(without alignments)
1321.448 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMTK.....KLMKRIAAKILSLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2720	100.0	525	5	US-10-665-990A-14	Sequence 14, Appl
2	2712	99.7	525	5	US-10-665-990A-16	Sequence 16, Appl
3	2708	99.6	525	5	US-10-665-990A-18	Sequence 18, Appl
4	2643	97.2	525	4	US-10-066-551-4	Sequence 4, Appli
5	2643	97.2	525	5	US-10-665-990A-4	Sequence 4, Appli
6	2643	97.2	525	5	US-10-665-990A-20	Sequence 20, Appl
7	632	23.2	502	4	US-10-335-977-9085	Sequence 9085, Ap
8	632	23.2	502	4	US-10-335-977-9086	Sequence 9086, Ap
9	590	21.7	428	3	US-09-861-752A-356	Sequence 356, App
10	359	13.2	224	4	US-10-335-977-9084	Sequence 9084, Ap
11	296	10.9	486	4	US-10-418-861B-59	Sequence 59, Appl
12	291	10.7	495	4	US-10-724-972A-3854	Sequence 3854, Ap
13	286.5	10.5	493	3	US-09-828-523A-22	Sequence 22, Appl
14	286.5	10.5	493	3	US-09-966-521-18	Sequence 18, Appl
15	286.5	10.5	493	4	US-10-429-094-18	Sequence 18, Appl
16	286.5	10.5	502	3	US-09-828-523A-90	Sequence 84, Appl
17	286.5	10.5	502	3	US-09-966-521-84	Sequence 84, Appl
18	286.5	10.5	502	4	US-10-429-094-84	Sequence 84, Appl
19	274.5	10.1	206	4	US-10-335-977-9083	Sequence 9083, Ap
20	259.5	9.5	494	4	US-10-724-972A-4379	Sequence 4379, Ap
21	252	9.3	422	4	US-10-627-476-78	Sequence 78, Appl
22	252	9.3	500	3	US-09-848-726-2	Sequence 2, Appli
23	252	9.3	500	3	US-09-738-626-6514	Sequence 6514, Ap
24	226	8.3	404	4	US-10-156-761-8603	Sequence 8603, Ap
25	219	8.1	510	4	US-10-474-776-225	Sequence 225, App
26	219	8.1	510	5	US-10-472-928-212	Sequence 212, App
27	215	7.9	518	5	US-10-617-320-3177	Sequence 3177, Ap

28	205	7.5	363	5	US-10-501-282-2384	Sequence 2384, Ap
29	205	7.5	448	5	US-10-501-282-2386	Sequence 2386, Ap
30	205	7.5	481	5	US-10-501-282-2388	Sequence 2388, Ap
31	205	7.5	493	5	US-10-501-282-2390	Sequence 2390, Ap
32	205	7.5	512	5	US-10-501-282-2392	Sequence 2392, Ap
33	175	6.4	280	5	US-10-450-763-31511	Sequence 31511, A
34	175	6.4	1097	5	US-10-450-763-31510	Sequence 31510, A
35	168.5	6.2	525	5	US-10-474-792-354	Sequence 354, App
36	141.5	5.2	476	4	US-10-312-273-75	Sequence 75, Appl
37	133.5	4.9	932	4	US-10-137-129A-6	Sequence 6, Appli
38	127	4.7	474	5	US-10-498-327-79	Sequence 79, Appl
39	124.5	4.6	548	4	US-10-408-765A-2203	Sequence 2203, Ap
40	122.5	4.5	553	4	US-10-424-599-215690	Sequence 215690, A
41	121.5	4.5	1683	4	US-10-369-493-22273	Sequence 22273, A
42	116.5	4.3	908	4	US-10-032-585-7642	Sequence 7642, Ap
43	116	4.3	851	4	US-10-437-963-120957	Sequence 120957, A
44	115.5	4.2	1046	4	US-10-437-963-175603	Sequence 175603, A
45	113.5	4.2	494	6	US-11-097-143-2316	Sequence 2316, Ap

ALIGNMENTS

RESULT 1
US-10-665-990A-14
; Sequence 14, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Nei
; TITLE OF INVENTION: infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; PRIOR FILING DATE: 2003-09-19
; PRIOR FILING DATE: 2003-07-15
; PRIOR FILING DATE: 2003-07-15
; PRIOR FILING DATE: 2002-01-31
; PRIOR FILING DATE: 2002-01-31
; PRIOR FILING DATE: 2001-10-23
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 2001-01-31
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae 1291
US-10-665-990A-14

Query Match	100.0%	Score	2720	DB	5	Length	525
Best Local Similarity	100.0%	Pred. No.	3.5e-233				
Matches	525	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
QY	1	MHTDPKIQAMPSETISPMTKRSLSLLCLLSCSSWLPPEERTSRHNTSKPVLLDN	60				
Db	1	MHTDPKIQAMPSETISPMTKRSLSLLCLLSCSSWLPPEERTSRHNTSKPVLLDN	60				
QY	61	ILQIRHTPHNGLSDIYLLDDPHEALAAARALTESAEHSLDLOYIYWRNDISGRLLFNLM	120				
Db	61	ILQIRHTPHNGLSDIYLLDDPHEALAAARALTESAEHSLDLOYIYWRNDISGRLLFNLM	120				
QY	121	YLAERGVRLLLDDNNTRGLDDLLALDSDHNPNI EVLFNPFVLRKRWALGYLTDPFRL	180				
Db	121	YLAERGVRLLLDDNNTRGLDDLLALDSDHNPNI EVLFNPFVLRKRWALGYLTDPFRL	180				
QY	181	NRMHNKSTADNRATILGGRNIGDGYFKVGEDTVFADLDILATGVSVGVSDFDORYA	240				
Db	181	NRMHNKSTADNRATILGGRNIGDGYFKVGEDTVFADLDILATGVSVGVSDFDORYA	240				

```
Qy 241 SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
|||||
Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
|||||
Qy 301 QTRLISDTPAKGLDRDRKPPPIAGRLQDALKEPEKSVYLVSPYFVPTKSGTDLAKLVQD 360
|||||
Db 301 QTRLISDTPAKGLDRDRKPPPIAGRLQDALKEPEKSVYLVSPYFVPTKSGTDLAKLVQD 360
|||||
Qy 361 GIDVTVLNSLOATDVAHVSGYVKYRKLKAGIKLYELOPNHAPATKDKGLTGSSVT 420
|||||
Db 361 GIDVTVLNSLOATDVAHVSGYVKYRKLKAGIKLYELOPNHAPATKDKGLTGSSVT 420
|||||
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQMERTLADTTPEYAYRV 480
|||||
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQMERTLADTTPEYAYRV 480
|||||
Qy 481 TLDKHNRLQWHDPATRKYTPNEPEAKLWKRIAANKILSLPIEGLL 525
|||||
Db 481 TLDKHNRLQWHDPATRKYTPNEPEAKLWKRIAANKILSLPIEGLL 525
|||||

RESULT 2
US-10-665-990A-16
; Sequence 16, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae 1090
US-10-665-990A-16

Query Match 99.7%; Score 2712; DB 5; Length 525;
Best Local Similarity 99.6%; Pred. No. 1.8e-232;
Matches 523; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPLLEERTESRHNFSKPVLND 60
|||||
Db 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPLLEERTESRHNFSKPVLND 60
|||||
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDLQYIWRNDISGRLLFNLM 120
|||||
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDLQYIWRNDISGRLLFNLM 120
|||||
Qy 121 YLAERGVVRVRLLLDDNNTGRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDPPRL 180
|||||
Db 121 YLAERGVVRVRLLLDDNNTGRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDPPRL 180
|||||
Qy 181 NRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVGVSHDFDRYWA 240
|||||
Db 181 NRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVGVSHDFDRYWA 240
|||||
Qy 241 SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
|||||
```

```
Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
|||||
Qy 301 QTRLISDTPAKGLDRDRKPPPIAGRLQDALKEPEKSVYLVSPYFVPTKSGTDLAKLVQD 360
|||||
Db 301 QTRLISDTPAKGLDRDRKPPPIAGRLQDALKEPEKSVYLVSPYFVPTKSGTDLAKLVQD 360
|||||
Qy 361 GIDVTVLNSLOATDVAHVSGYVKYRKLKAGIKLYELOPNHAPATKDKGLTGSSVT 420
|||||
Db 361 GIDVTVLNSLOATDVAHVSGYVKYRKLKAGIKLYELOPNHAPATKDKGLTGSSVT 420
|||||
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQMERTLADTTPEYAYRV 480
|||||
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQMERTLADTTPEYAYRV 480
|||||
Qy 481 TLDKHNRLQWHDPATRKYTPNEPEAKLWKRIAANKILSLPIEGLL 525
|||||
Db 481 TLDKHNRLQWHDPATRKYTPNEPEAKLWKRIAANKILSLPIEGLL 525
|||||

RESULT 3
US-10-665-990A-18
; Sequence 18, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae ms11
US-10-665-990A-18

Query Match 99.6%; Score 2708; DB 5; Length 525;
Best Local Similarity 99.4%; Pred. No. 4.1e-232;
Matches 522; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPLLEERTESRHNFSKPVLND 60
|||||
Db 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPLLEERTESRHNFSKPVLND 60
|||||
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDLQYIWRNDISGRLLFNLM 120
|||||
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDLQYIWRNDISGRLLFNLM 120
|||||
Qy 121 YLAERGVVRVRLLLDDNNTGRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDPPRL 180
|||||
Db 121 YLAERGVVRVRLLLDDNNTGRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDPPRL 180
|||||
Qy 181 NRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVGVSHDFDRYWA 240
|||||
Db 181 NRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVGVSHDFDRYWA 240
|||||
Qy 241 SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
|||||
Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
|||||
```

Qy 301 QTRLISDTPAKGLDRDRRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Qy 361 GIDVTVLTNSLOATDVAHVSGYVKYRKPILLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLTNSLOATDVAHVSGYVKYRKPILLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRFIFGFSNLDPR SARLNTMGVVIESP KIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRFIFGFSNLDPR SARLNTMGVVIESP KIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHD PATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TLDKHNRLQWHD PATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525

RESULT 4

US-10-066-551-4
; Sequence 4, Application US/10066551
; Publication No. US20030100071A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, M. A.
; APPLICANT: Edwards, J. L.
; APPLICANT: Gibson, B. W.
; APPLICANT: Scheffler, K.
; APPLICANT: Brown, E.
; TITLE OF INVENTION: Vaccine and compositions for the
; prevention and treatment of Neisserial infections
; FILE REFERENCE: 875.045U1
; CURRENT APPLICATION NUMBER: US/10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 4
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-066-551-4

Query Match 97.2%; Score 2643; DB 4; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.6e-226;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPLERTESRHFNTSKPVLDDN 60
Db 1 MRANPKTQAMPSETISLMKTRSLISLLCLLSCSSWLPPLERTESRHFNTSKPVLDDN 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLPNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAFAARAALIESAHSLLDQYIWRNDISGRLLPNLV 120
Qy 121 YLAAAGRVRLLLDDNNTGRLDILLALDHPNIEVRLNPNFVLRKWRALGYLTDPPRL 180
Db 121 YLAAAGRVRLLLDDNNTGRLDILLALDHPNIEVRLNPNFVLRKWRALGYLTDPPRL 180
Qy 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGVSHDFDRYWA 240
Db 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGVSHDFDRYWA 240
Qy 241 SHSAHNATRIIRSGNIKGLOALGYNDETSRHALLRVRETVESPLQYKIQTRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIKGLOALGYNDETSRHALLRVRETVESPLQYKIQTRIDWQSV 300
Qy 301 QTRLISDTPAKGLDRDRRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360

Qy 361 GIDVTVLTNSLOATDVAHVSGYVKYRKPILLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLTNSLOATDVAHVSGYVKYRKPILLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRFIFGFSNLDPR SARLNTMGVVIESP KIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRFIFGFSNLDPR SARLNTMGVVIESP KIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHD PATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TLDKHNRLQWHD PATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525

RESULT 5

US-10-665-990A-4
; Sequence 4, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Nei
; TITLE OF INVENTION: infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; PRIOR FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-665-990A-4

Query Match 97.2%; Score 2643; DB 5; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.6e-226;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPLERTESRHFNTSKPVLDDN 60
Db 1 MRANPKTQAMPSETISLMKTRSLISLLCLLSCSSWLPPLERTESRHFNTSKPVLDDN 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLPNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAFAARAALIESAHSLLDQYIWRNDISGRLLPNLV 120
Qy 121 YLAAAGRVRLLLDDNNTGRLDILLALDHPNIEVRLNPNFVLRKWRALGYLTDPPRL 180
Db 121 YLAAAGRVRLLLDDNNTGRLDILLALDHPNIEVRLNPNFVLRKWRALGYLTDPPRL 180
Qy 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGVSHDFDRYWA 240
Db 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGVSHDFDRYWA 240
Qy 241 SHSAHNATRIIRSGNIKGLOALGYNDETSRHALLRVRETVESPLQYKIQTRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIKGLOALGYNDETSRHALLRVRETVESPLQYKIQTRIDWQSV 300
Qy 301 QTRLISDTPAKGLDRDRRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Qy 361 GIDVTVLTNSLOATDVAHVSGYVKYRKPILLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLTNSLOATDVAHVSGYVKYRKPILLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420

Db 361 GIDVTVLNSLOATDVAHVHGVYKVRKPLKAGIKLYELOPNHVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525

RESULT 6
US-10-665-990A-20
; Sequence 20, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neisseria meningitidis B
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; PRIOR FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; TYPE: PRT
; ORGANISM: Neisseria meningitidis B
US-10-665-990A-20

Query Match 97.2%; Score 2643; DB 5; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.6e-226;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MHTDPTQAMPSETISPMKTRSLISLLCLLCCSSWLPPLLEERTESRHFNTSKPVLND 60
Db 1 MRANPKTQAMPSETISLMKTRSLISLLCLLCCSSWLPPLLEERTESRHFNTSKPVLND 60

Qy 61 ILQIRHTPHNGLSDIYLLDDPHEALAAARALIESAEHSLDLQYIWRNDISGRLLFNLM 120
Db 61 ILQIRHTPHNGLSDIYLLNDPHEAFARAALIESAEHSLDLQYIWRNDISGRLLFNLM 120

Qy 121 YLAERGVVRVLLDDNNTRGLDILLALDHPNIEVLFNPFVLRKWRALGYLTFPRL 180
Db 121 YLAERGVVRVLLDDNNTRGLDILLALDHPNIEVLFNPFVLRKWRALGYLTFPRL 180

Qy 181 NRRMKNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDFRYWA 240
Db 181 NRRMKNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDFRYWA 240

Qy 241 SHSAINATRIIRSGNIGKGLQALQYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
Db 241 SHSAINATRIIRSGNIGKGLQALQYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSV 300

Qy 301 QTRLSDTPAKGLDRDRKPPITAGRLQDALKOPEKSVVLVSPYFVPTKSGTDLAKLVQD 360
Db 301 QTRLSDTPAKGLDRDRKPPITAGRLQDALKOPEKSVVLVSPYFVPTKSGTDLAKLVQD 360

Qy 361 GIDVTVLNSLOATDVAHVHGVYKVRKPLKAGIKLYELOPNHVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLOATDVAHVHGVYKVRKPLKAGIKLYELOPNHVPATKDKGLTGSSVT 420

Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525

RESULT 7
US-10-335-977-9085
; Sequence 9085, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO HELICOBACTER PYLORI FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9085:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...502
; SEQUENCE DESCRIPTION: SEQ ID NO: 9085:
US-10-335-977-9085

Query Match 23.2%; Score 632; DB 4; Length 502;
Best Local Similarity 32.8%; Pred. No. 4e-47;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;

Qy 64 IRHTPHNGLSDIY-----LDDPHEALAAARALIESAEHSLDLQYIWRN 109
Db 30 ISYDPYTTTIGSLYAKNLKENPKHSAAILLEDGFDALLHRVGLIRMSQKSIDMQTYIKN 89

Qy 110 DISGELLENMLAERGVVRVLLDDNNTRGLD----DLLALDSDHPNIEVLRNPFVL 165
Db 90 DLSQVITAKELNLNANRGVVRILLDDN---GLSDPFDIML-LNFHKNIEVKIENFYI 145

Qy 166 RKWRALGY---LTDFFPRLNRRMKNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDIL 222

Db 146 RN-KGLRYFEMLADYERIKRMHNLFIYDVFVAVIIGRNIGDNYFNDLDTNFDLIDLAL 204
Qy 223 ATGSVVGVSVDHFDYRWASHSAHNATRIIRS-----GNIGKGLQALGVNDETSRHAL 274
Db 205 FFGVASKAKESFENYRFRHSIPVS--LLRTHKLNKNNVKEIAKLHEKIPISAEADANEFE 263
Qy 275 LRYRETVEQSPLYQ-KIQTGRIDWQSVQTRLSIDTPAKGLDRDRKPPPIAGRLQDALKQP 333
Db 264 KKVNDFIERFQKYQPIYVG-----NAIFLADLPK-IDTPLYS-PIKIAFEKALKNA 314
Qy 334 EKSIVLVSPYFVPTKSGTDALAKLVQDIDVTVLNLSQATDVAHVHGVYKRPKLKA 393
Db 315 KDSVFIASSYFIPGKIMKIFKNQISKGIELNLTNSLSDAIVVYGAWERYNKLVRM 374
Qy 394 GIKLYELQPNHVAVPATKDKGLTGSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNTMG 453
Db 375 GANVYEIRNDFNFRQIKGR----FTKLSLHGKTVFDDALTLGSPNIDPRSAINTESA 431
Qy 454 VVIESPKIAEQMERTLADTTPEYAYRVTLDKHNLQWHDPAKTRKTY----PNEPEAKLWK 509
Db 432 VLFDPNPSFAKRVRLSLKDHA-QQSWHLVLYRH-RVWE--ATEBGILIHKNSPDTSFFL 487
Qy 510 RIAAKILSLLP 520
Db 488 RLKIEWSKVLP 498
RESULT 8
US-10-335-977-9086
; Sequence 9086, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (B) LOCATION 1...502
; SEQUENCE DESCRIPTION: SEQ ID NO: 9086:
US-10-335-977-9086
Query Match 23.2%; Score 632; DB 4; Length 502;
Best Local Similarity 32.8%; Pred. No. 4e-47;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;
Qy 64 IRHTPHNNGSLDIY-----LLDDPHEALAAARALIESAEHSLDLOYIWRN 109
Db 30 ISVDYPTTTTIGSYAKNLKENPKHSAAILLEDGFDALLHRVGLIRMSQKSIDMQTYIKN 89
Qy 110 DISGRILFNLMYLAERGVRLLLDNNTRGLD---DLLLALDSHPNIEVRLFPVL 165
Db 90 DLSQVIATKELLNANRGVVRILLDDN---GLDSDPSDML-LNFHKNIEVKIFNFIYI 145
Qy 166 RKWRALGY---LTFPRLNRRMNKSGFTADNRATILGGRNIGDEYFKVGEDTVFADLDIL 222
Db 146 RN-KGLRYFEMLADYERIKRMHNLFIYDVFVAVIIGRNIGDNYFNDLDTNFDLIDLAL 204
Qy 223 ATGSVVGVSVDHFDYRWASHSAHNATRIIRS-----GNIGKGLQALGVNDETSRHAL 274
Db 205 FFGVASKAKESFENYRFRHSIPVS--LLRTHKLNKNNVKEIAKLHEKIPISAEADANEFE 263
Qy 275 LRYRETVEQSPLYQ-KIQTGRIDWQSVQTRLSIDTPAKGLDRDRKPPPIAGRLQDALKQP 333
Db 264 KKVNDFIERFQKYQPIYVG-----NAIFLADLPK-IDTPLYS-PIKIAFEKALKNA 314
Qy 334 EKSIVLVSPYFVPTKSGTDALAKLVQDIDVTVLNLSQATDVAHVHGVYKRPKLKA 393
Db 315 KDSVFIASSYFIPGKIMKIFKNQISKGIELNLTNSLSDAIVVYGAWERYNKLVRM 374
Qy 394 GIKLYELQPNHVAVPATKDKGLTGSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNTMG 453
Db 375 GANVYEIRNDFNFRQIKGR----FTKLSLHGKTVFDDALTLGSPNIDPRSAINTESA 431
Qy 454 VVIESPKIAEQMERTLADTTPEYAYRVTLDKHNLQWHDPAKTRKTY----PNEPEAKLWK 509
Db 432 VLFDPNPSFAKRVRLSLKDHA-QQSWHLVLYRH-RVWE--ATEBGILIHKNSPDTSFFL 487
Qy 510 RIAAKILSLLP 520
Db 488 RLKIEWSKVLP 498
RESULT 9
US-09-881-752A-356
; Sequence 356, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1e1 Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-356
Query Match 21.7%; Score 590; DB 3; Length 428;
Best Local Similarity 32.3%; Pred. No. 1.8e-43;


```
Qy 376 VAAHSGVVKYRKPILLKAGIKLYELQPNHVPATKDKLGTSSVTSLSHAKTFIVDGKRIF 435
Db 381 HPLVYWATFSNASDLSGSKYI-----TYENG-----IHSKMLDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFVVDENLAKDL-----RVAYEHDTKSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFESLAKLVS 490

RESULT 14
US-09-966-521-18
; Sequence 18, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 6212.N2
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-966-521-18

Query Match 10.5%; Score 286.5; DB 3; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYDQVLKD-----IKN 151
Qy 96 AEHSLDLOYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPNI 155
Db 152 AKEVILHEYYTFALDGLGKRILHALEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVRKRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGRTNGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLLNFRMNNHRKIIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGVRDTHLRITQGDVAVALQRLFILDWNS-QAHRPQFEYDVKKYPPKNGPLG- 311
Qy 259 GLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQ---TFLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMIMSA----- 340
Qy 316 DRRKPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLNLSQATD 375
Db 341 -----KKSVMYLOSYPFI PDNSYINAIKIAAKSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYRKPILLKAGIKLYELQPNHVPATKDKLGTSSVTSLSHAKTFIVDGKRIF 435
Db 381 HPLVYWATFSNASDLSGSKYI-----TYENG-----IHSKMLDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFVVDENLAKDL-----RVAYEHDTKSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFESLAKLVS 490

RESULT 15
US-10-429-094-18
; Sequence 18, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774.US1 CNI
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-429-094-18

Query Match 10.5%; Score 286.5; DB 4; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYDQVLKD-----IKN 151
Qy 96 AEHSLDLOYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPNI 155
Db 152 AKEVILHEYYTFALDGLGKRILHALEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVRKRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGRTNGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLLNFRMNNHRKIIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGVRDTHLRITQGDVAVALQRLFILDWNS-QAHRPQFEYDVKKYPPKNGPLG- 311
Qy 259 GLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQ---TFLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMIMSA----- 340
Qy 316 DRRKPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLNLSQATD 375
Db 341 -----KKSVMYLOSYPFI PDNSYINAIKIAAKSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYRKPILLKAGIKLYELQPNHVPATKDKLGTSSVTSLSHAKTFIVDGKRIF 435
Db 381 HPLVYWATFSNASDLSGSKYI-----TYENG-----IHSKMLDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFVVDENLAKDL-----RVAYEHDTKSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFESLAKLVS 490

RESULT 16
US-09-828-523A-90
; Sequence 90, Application US/09828523A
; Patent No. US20020168697A1
; GENERAL INFORMATION:
```

```
Db 469 KESYANRPLSVKFESLAKLVS 490

RESULT 15
US-10-429-094-18
; Sequence 18, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774.US1 CNI
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-429-094-18

Query Match 10.5%; Score 286.5; DB 4; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYDQVLKD-----IKN 151
Qy 96 AEHSLDLOYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPNI 155
Db 152 AKEVILHEYYTFALDGLGKRILHALEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVRKRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGRTNGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLLNFRMNNHRKIIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGVRDTHLRITQGDVAVALQRLFILDWNS-QAHRPQFEYDVKKYPPKNGPLG- 311
Qy 259 GLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQ---TFLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMIMSA----- 340
Qy 316 DRRKPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLNLSQATD 375
Db 341 -----KKSVMYLOSYPFI PDNSYINAIKIAAKSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYRKPILLKAGIKLYELQPNHVPATKDKLGTSSVTSLSHAKTFIVDGKRIF 435
Db 381 HPLVYWATFSNASDLSGSKYI-----TYENG-----IHSKMLDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFVVDENLAKDL-----RVAYEHDTKSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFESLAKLVS 490

RESULT 16
US-09-828-523A-90
; Sequence 90, Application US/09828523A
; Patent No. US20020168697A1
; GENERAL INFORMATION:
```

; APPLICANT: The Pharmacia & Upjohn Company
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 268.62120101
; CURRENT APPLICATION NUMBER: US/09/828,523A
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/266,327
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned for
; OTHER INFORMATION: expression in E. coli.
US-09-828-523A-90

Query Match 10.5%; Score 286.5; DB 3; Length 502;
Best Local Similarity 21.1%; Pred. No. 2.5e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;
QY 42 EERTESRHFTNSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAALIES 95
DB 105 DNKOVQKHDLVRLLMDQDGFLENKVDHFDIGNDLYDQVLKD-----IKN 152
QY 96 AEHSLDLYIWRNDISGRLLFNILMYLAAERGVVRLLDDNNTRGLDLDLLALDHPNI 155
DB 153 AKEVIHLEYTFALDGLGKILHALEKLGKLEKLYDDVGSK-----NV 199
QY 156 EVRLNFPVLRKWRALG-----YLTDFP-----RLNRRMHNKSTADNRATILGGRNIGD 205
DB 200 KMANFDHF-----KSLGGEVEAFASKPLPLNFRMNRNRKIIIVIDGQLGVYGVGNIGD 254
QY 206 EYFKVGEDTVFADLDILATGCVGEVSHDPRYWAHSAHNAIRII-----RSGNIGK 258
DB 255 EYLGKLGKGYWRDTHLRIQGDALQRLFDLWNS-QAHRPOFEYDVKYFPKKNGLPLG- 312
QY 259 GLQALGYNDTSHALLRYRETVESQSPLYQKIOTGRIDWOSVQ---TRLISDTPAKGLDR 315
DB 313 -----NSPIQTAASGPASDWHQIEGYTKMMSA----- 341
QY 316 DRRKPPIAGRLQDALQKPEKSVLYSPYFPTKSGTDALAKLVODGIDVTVLNLSQATD 375
DB 342 -----KKSIVLQSPYFIPDINSYINAIKIAAKSGVDVHLMIPC--KPD 381
QY 376 VAAVHSGYKVRKPLKAGIKLYELQPNHAVPATKDKLTGSSVTSLSHAKTIFVDGKRIF 435
DB 382 HPLVYWATFNSADLLSSGVKIY-----TYENG-----IHSKMCLIDDEIVS 424
QY 436 IGSFNLDPRGARLNTMGVVIESPKEAEOEMERTLADTTPYAVRVTLDKHNRQLQWHPAT 495
DB 425 VGTANMDFRSFELNFEVNAFYDENLAKDL-----RVAYEHDITKSKQL-----T 469
QY 496 RKTYPNEPEAKLMKRIAAKILS 517
DB 470 KESYANRPLSVKPKESLAKLVS 491

RESULT 17
US-09-966-521-84
; Sequence 84, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 6212.N2
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28

; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned for
; OTHER INFORMATION: expression in E. coli
US-09-966-521-84

Query Match 10.5%; Score 286.5; DB 3; Length 502;
Best Local Similarity 21.1%; Pred. No. 2.5e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;
QY 42 EERTESRHFTNSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAALIES 95
DB 105 DNKOVQKHDLVRLLMDQDGFLENKVDHFDIGNDLYDQVLKD-----IKN 152
QY 96 AEHSLDLYIWRNDISGRLLFNILMYLAAERGVVRLLDDNNTRGLDLDLLALDHPNI 155
DB 153 AKEVIHLEYTFALDGLGKILHALEKLGKLEKLYDDVGSK-----NV 199
QY 156 EVRLNFPVLRKWRALG-----YLTDFP-----RLNRRMHNKSTADNRATILGGRNIGD 205
DB 200 KMANFDHF-----KSLGGEVEAFASKPLPLNFRMNRNRKIIIVIDGQLGVYGVGNIGD 254
QY 206 EYFKVGEDTVFADLDILATGCVGEVSHDPRYWAHSAHNAIRII-----RSGNIGK 258
DB 255 EYLGKLGKGYWRDTHLRIQGDALQRLFDLWNS-QAHRPOFEYDVKYFPKKNGLPLG- 312
QY 259 GLQALGYNDTSHALLRYRETVESQSPLYQKIOTGRIDWOSVQ---TRLISDTPAKGLDR 315
DB 313 -----NSPIQTAASGPASDWHQIEGYTKMMSA----- 341
QY 316 DRRKPPIAGRLQDALQKPEKSVLYSPYFPTKSGTDALAKLVODGIDVTVLNLSQATD 375
DB 342 -----KKSIVLQSPYFIPDINSYINAIKIAAKSGVDVHLMIPC--KPD 381
QY 376 VAAVHSGYKVRKPLKAGIKLYELQPNHAVPATKDKLTGSSVTSLSHAKTIFVDGKRIF 435
DB 382 HPLVYWATFNSADLLSSGVKIY-----TYENG-----IHSKMCLIDDEIVS 424
QY 436 IGSFNLDPRGARLNTMGVVIESPKEAEOEMERTLADTTPYAVRVTLDKHNRQLQWHPAT 495
DB 425 VGTANMDFRSFELNFEVNAFYDENLAKDL-----RVAYEHDITKSKQL-----T 469
QY 496 RKTYPNEPEAKLMKRIAAKILS 517
DB 470 KESYANRPLSVKPKESLAKLVS 491

RESULT 18
US-10-429-094-84
; Sequence 84, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774.US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 502
; TYPE: PRT

Best Local Similarity 22.3%; Pred. No. 6.2e-14;
Matches 101; Conservative 70; Mismatches 178; Indels 103; Gaps 11;
Qy 76 IYLLDDPHEALAAARAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVRLD 135
Db 133 IDLFTDGHKLVKEVDIYNAQDIHLEYTFELDGLGKILDALETKLEGLEVLUYD 192
Qy 136 DNNTRGLDLLLLLALDHPNIEVLPFVLRKWRALGYLTD--PP-----RLNRRMH 185
Db 193 DVGSK-----KVLRSK---FKHFRALGGEVEAPPPSKVPLINFRMNRNH 234
Qy 186 NKSTADNRATILGRNIGDEYFKVGEDTVPADLDILATGSVGEVSHDFDRYWASHAH 245
Db 235 RKIIIDGQIGYGVGFNGVDYGLGKLGWYRDTHTRVQECIDALQLRFILWNSSQ-- 292
Qy 246 NATRIIRSGNIGKLOALGYNDETSHALLRYRTEVQSPLYQKIOTGRIDWQSVQRLI 305
Db 293 -----HRPQFKDQYFPKKG--DKGNAALQIA 319
Qy 306 SDTPAKGLDRRRKPPIAGRLQDALKOPEKSVYLVSPYFTKSGTDALAKLVQDGIDVT 365
Db 320 SSGPAFLHQ-----IEYGYTKMMSAKKSIYLSQSYFIPDQSYINALKWAANSQVEVN 373
Qy 366 VLTNSLOATVAAVHSGYVYKRPGLLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAK 425
Db 374 LMIPC--KPHPFYVWATFNAADLLDSGVNIYTYQNGF-----IHSK 414
Qy 426 TFIVDGKRIFIGSNLDPRLNTEMGVNVIESPKIAEQMERTLADTTPEYAVRVTLDKH 485
Db 415 ILMIDDEISSGNMDFRFLNFVNAFYBEDIAKLRQ-----AFKEDIEQS 465
Qy 486 NRLQWHDPAKRTKYPNPEAKLWKRIAAKILS 517
Db 466 KLL-----TKEVYDKRPLSIKFKEGLAKLIS 491

RESULT 21

US-10-627-476-78
; Sequence 78, Application US/10627476
; Publication No. US20040030116A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schoder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-125CPCN
; CURRENT APPLICATION NUMBER: US/10/627,476
; CURRENT FILING DATE: 2003-07-25
; PRIOR FILING DATE: 09/602,787
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 78
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-627-476-78

Query Match 9.3%; Score 252; DB 4; Length 422;
Best Local Similarity 23.4%; Pred. No. 2.3e-13;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;
Qy 67 TPNNGLSDIYLLDDPHEALAAARAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAAR 126
Db 38 TGGNNG-----FYSDYRESLKRTAAIDEAEVEIYVEIYIMAWSYTQPPFAALERAHNR 92
Qy 127 GVRVRLD-----DNNTRGLDLLLLLALDHPNIEVLPFVLRKWRAL 171
Db 93 GVKVRLLDHVGWVKPYGHRLLKELNRMGFAYLM-----LPLQWR-- 135
Qy 172 GYLTDPPRLNRRMHNKSFTADNRATILGGRN-IGDEY-----PKVGBDTPVADLDILATG 225
Db 136 ---RRFRPDLNRHKKWLIIDGHTAPMGSQLIAPSYLOKKNIKLGRE--WKDLMLVELTG 190
Qy 226 SVVGEVSHDFDRYWASHAHNATRIIRSGNIGKLOALGYNDETSHALLRYRTEVQS 285
Db 191 PIVSSMEMIPAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 234
Qy 286 LYQKIOTGRIDWQSVQTRLISDTPAKGLDRDRKPPIAGRLQDALKOPEKSVYLVSPFV 345
Db 235 LVQLIPSG-----PGYTTEPNLR-----MFSIVHHAKERLILCSPYFI 273
Qy 346 PTKSGTDALAKLVQDGIDVTLTNSLOATVAAVHSGYVYKRPGLLKAGIKLYELQPNHA 405
Db 274 PDESLEAVTSACYRGVTVELFVS--EQADQFAIDHAQSSYYQALLEAGYKIQF-PKPD 330
Qy 406 VPATKDKGLTGSSVTSLSHAKTFIVD-----GKRIF--IGSFNLDPPRSARLNTMGVVIES 458
Db 331 V-----LHTKYMIAADPDITGNEALGVLGSSNLDIRSFGLNYSLSLMIK 375
Qy 459 PKIAEQMERTLADTTPEYAVRVTLDKHNRLOWHDPATRTKY 499
Db 376 GNLIHEL-NALTDRYRTVSFKLTLDKWNQRSW-----RRRY 410

RESULT 22

US-09-848-726-2
; Sequence 2, Application US/09848726
; Patent No. US20020102667A1
; GENERAL INFORMATION:
; APPLICANT: NAMPOOTHIRI, Madhavan
; TITLE OF INVENTION: No. US20020102667A1el Nucleotide Sequences Coding for the cls
; FILE REFERENCE: 032301 WD 1191
; CURRENT APPLICATION NUMBER: US/09/848,726
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-848-726-2

Query Match 9.3%; Score 252; DB 3; Length 500;
Best Local Similarity 23.4%; Pred. No. 3e-13;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;
Qy 67 TPNNGLSDIYLLDDPHEALAAARAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAAR 126
Db 116 TGGNNG-----FYSDYRESLKRTAAIDEAEVEIYVEIYIMAWSYTQPPFAALERAHNR 170
Qy 127 GVRVRLD-----DNNTRGLDLLLLLALDHPNIEVLPFVLRKWRAL 171
Db 171 GVKVRLLDHVGWVKPYGHRLLKELNRMGFAYLM-----LPLQWR-- 213

QY 172 GYLTFPRLNRRMHNKSFADNRATILGGRN-IGDEY-----FKVGEDTVFADLDILATG 225
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 214 ---RRFRPDLNRHRKMLIIDGHTAFMGSQNLIAPSYQKKNIKLGRB--WKDLMLVELTG 268
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 226 SVVGEVSHDFRYWASHSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRVRETVEOSP 285
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 269 PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 312
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 286 LYQKITQGRIDWQSVQTRLSIDTPAKGLDRDRRKPPIAGRIQDALKQPEKSVYLVSPIYV 345
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 269 PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 312
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 286 LYQKITQGRIDWQSVQTRLSIDTPAKGLDRDRRKPPIAGRIQDALKQPEKSVYLVSPIYV 345
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 313 LVQLIPSG-----PGYTPEPNLR-----MENSIVHHAKEKRLILCSPYFI 351
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 346 PTKSGTDALAKLVQGDIDVTVLNLSQATDVAAVHSGVYKVRKPLLKAGIKLYELQPNHA 405
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 313 LVQLIPSG-----PGYTPEPNLR-----MENSIVHHAKEKRLILCSPYFI 351
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 346 PTKSGTDALAKLVQGDIDVTVLNLSQATDVAAVHSGVYKVRKPLLKAGIKLYELQPNHA 405
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 352 PDESLEAVTSACRYGVTVLFSV--EQADQFAIDHAQSSYYQALLEAGVKIYQF-PKPD 408
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 406 VPATKDKGLTSSVTSLSHAKTFIVD-----GKRIF--IGSNLDPRSARLNTMGVWIES 458
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 409 V-----LHTKYMADPDDTTGNEALGVLGSSNLDIRSFGNLVEISLMIK 453
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 459 PKIAEQMERTLADTTPPEYAVYRTLDKHNRLQWHPDPAKTKY 499
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 454 GNLIHEL-NALTDYRTVTSFKLTLCKMQRNSW-----RRRY 488

RESULT 23

US-09-738-626-6514
; Sequence 6514, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6514
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6514

Query Match 9.3%; Score 252; DB 3; Length 500;
Best Local Similarity 23.4%; Pred. No. 3e-13;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;

QY 67 THNNGLSDIYLLDDPHEALARAALIESAHSLDLQYIWRNDISGRLLNMLYLAER 126
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 116 TGGNNG-----FYSYRESLKMATAIDEABEYIYVEIYIMAWDSYTOPFFAALERAHNR 170
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 127 GVRVRLLED-----DNTRGLDLDLALDSHPNIEVRLNFPVLRKWRAL 171
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 171 GVKVRLLEDHVGSKYKPGYHRLKKELNRMGFANYLM-----LPLQPR-- 213
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 172 GYLTFPRLNRRMHNKSFADNRATILGGRN-IGDEY-----FKVGEDTVFADLDILATG 225
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Db 214 ---RRFRPDLNRHRKMLIIDGHTAFMGSQNLIAPSYQKKNIKLGRB--WKDLMLVELTG 268
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 226 SVVGEVSHDFRYWASHSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRVRETVEOSP 285
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 269 PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 312
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 286 LYQKITQGRIDWQSVQTRLSIDTPAKGLDRDRRKPPIAGRIQDALKQPEKSVYLVSPIYV 345
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 313 LVQLIPSG-----PGYTPEPNLR-----MENSIVHHAKEKRLILCSPYFI 351
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 346 PTKSGTDALAKLVQGDIDVTVLNLSQATDVAAVHSGVYKVRKPLLKAGIKLYELQPNHA 405
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 352 PDESLEAVTSACRYGVTVLFSV--EQADQFAIDHAQSSYYQALLEAGVKIYQF-PKPD 408
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 406 VPATKDKGLTSSVTSLSHAKTFIVD-----GKRIF--IGSNLDPRSARLNTMGVWIES 458
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 409 V-----LHTKYMADPDDTTGNEALGVLGSSNLDIRSFGNLVEISLMIK 453
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 459 PKIAEQMERTLADTTPPEYAVYRTLDKHNRLQWHPDPAKTKY 499
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 454 GNLIHEL-NALTDYRTVTSFKLTLCKMQRNSW-----RRRY 488

RESULT 24

US-10-156-761-8603
; Sequence 8603, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8603
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8603

Query Match 8.3%; Score 226; DB 4; Length 404;
Best Local Similarity 23.8%; Pred. No. 4.4e-11;
Matches 100; Conservative 55; Mismatches 170; Indels 96; Gaps 16;

QY 58 LDNLIQIRHTPHNNGLSDIYLLDDPHEALARAALIESAHSLDLQYI--WRNDISGRLL 116
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 31 LERLIGIAATEGN-----ELVPLRNGDEIFPAMLRAIRTAETHVDMTFTVYWR----GGIA 82
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 117 FNLMYLAER---GVRVRLDDNNTRGLDLDLALDSHPNIEVRLNFPVLRKWRALGY 173
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 83 RDPAAALAEARAGVRVRLLDGFGAKQIEBGLLESMSAGVEVAMF-----RKPLWLSP 137
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 174 LTDPLNRRMHNKSFADNRATILGGRNIGDEYFKVGED-----TVFADLDILATGSVVG 229
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 138 L-----KQNRCHRKVLYVDEHTAFTGGVGVIAEAW---GGDARNPSEWEDTHVQVRGPAVD 190
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 230 EVSHDFRYWASHSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRVRETVEQS--PLY 287
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 191 GIAAFAQNAW--ECHD-----ELFDERDRFTEH-----EQSGRATV 225
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 288 QKIQ-TGRIDWQSVQTRLSIDTPAKGLDRDRRKPPIAGRIQDALKQPEKSVYLVSPIYV 346
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 226 QWVRGSASFGWDMQT-----LVRVMLASAEERFRLSTAYFAP 263
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2002-11-18
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2386
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Alloioococcus otitidis
US-10-501-282-2386

Query Match 7.5%; Score 205; DB 5; Length 448;
Best Local Similarity 21.1%; Pred. No. 3.9e-09;
Matches 92; Conservative 64; Mismatches 151; Indels 130; Gaps 17;

QY	80	DDPHEALAAARAALIESAEHSLDLOYYIWRNDISGRLLFNLMYLAARGVVRLLDD--NN	138
Db	90	EDLFEALKED---LRQAKYIFLEFYILEKGTMLDEILELFDKAEGLDIRLLYDDVGN	146
QY	139	TRGLDDLLALDGHNPNEVRLFPVLRKWRALGYLTFDPRLRNRHMKSFADNRATIL	198
Db	147	ILRIDENFVSYLENEKIKTSVFNPL---DWR---LTF---QYNYVRDRKIMVVDGKIGYT	197
QY	199	GGRNIGDEYF---KVGE-----DTVFADLDILATGSSVVGVEVSHDF---	235
Db	198	GGMNIGDNYNRVEKAGHWKGGIRLEGQGVWGFTTMF-----LSLWDYLNDDTDQDFRDF	252
QY	236	--DRYWASHAHNATRIIRSGNIGKGLQALGY-----NDETSRHALLRYRETVEQSPLYQ	288
Db	253	HPDEY-----KSGANEVVTGP-----GYVHPFADDPNTNRIQI-----SESLYL	289
QY	289	KIOTGRIDMQSVQTRLIISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPFYFPTK	348
Db	290	KL-----IYNAKESIYLVKTPYLFISQ	310
QY	349	SGTDALAKLVQDIDVTVLNSLOATDVAAGVHSGYVKRKPCLKAGIKLYELOPNHAPVA	408
Db	311	KLYSALENAALSGVDVRIVTPGI--PDKKIVFETTQSFYDKLLEVGKVIYVAPGF---	364
QY	409	TKDKGLTGSSTVSLHAKTFIVDGKRFIGSFNLDPRSARLNTMGVVIESPKEAQWERT	468
Db	365	-----IHEKVIIDQDFAINGTINFYRSLHHSFECGVLFYNTQSIIDMKND	411
QY	469	LADTTPAYRYVRTLDKH	485
Db	412	FDNLFP-ICRQVSLEEN	427

RESULT 30
US-10-501-282-2386
; Sequence 2386, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2

; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2002-11-18
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2388
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Alloioococcus otitidis
US-10-501-282-2388

QY	80	DDPHEALAAARAALIESAEHSLDLOYYIWRNDISGRLLFNLMYLAARGVVRLLDD--NN	138
Db	123	EDLFEALKED---LRQAKYIFLEFYILEKGTMLDEILELFDKAEGLDIRLLYDDVGN	179
QY	139	TRGLDDLLALDGHNPNEVRLFPVLRKWRALGYLTFDPRLRNRHMKSFADNRATIL	198
Db	180	ILRIDENFVSYLENEKIKTSVFNPL---DWR---LTF---QYNYVRDRKIMVVDGKIGYT	230
QY	199	GGRNIGDEYF---KVGE-----DTVFADLDILATGSSVVGVEVSHDF---	235
Db	231	GGMNIGDNYNRVEKAGHWKGGIRLEGQGVWGFTTMF-----LSLWDYLNDDTDQDFRDF	285
QY	236	--DRYWASHAHNATRIIRSGNIGKGLQALGY-----NDETSRHALLRYRETVEQSPLYQ	288
Db	286	HPDEY-----KSGANEVVTGP-----GYVHPFADDPNTNRIQI-----SESLYL	322
QY	289	KIOTGRIDMQSVQTRLIISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPFYFPTK	348
Db	323	KL-----IYNAKESIYLVKTPYLFISQ	343
QY	349	SGTDALAKLVQDIDVTVLNSLOATDVAAGVHSGYVKRKPCLKAGIKLYELOPNHAPVA	408
Db	344	KLYSALENAALSGVDVRIVTPGI--PDKKIVFETTQSFYDKLLEVGKVIYVAPGF---	397
QY	409	TKDKGLTGSSTVSLHAKTFIVDGKRFIGSFNLDPRSARLNTMGVVIESPKEAQWERT	468
Db	398	-----IHEKVIIDQDFAINGTINFYRSLHHSFECGVLFYNTQSIIDMKND	444
QY	469	LADTTPAYRYVRTLDKH	485
Db	445	FDNLFP-ICRQVSLEEN	460

RESULT 31
US-10-501-282-2390
; Sequence 2390, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2002-11-18
; PRIOR FILING DATE: 2002-11-25


```
Best Local Similarity 20.4%; Pred. No. 8.9e-07;
Matches 83; Conservative 44; Mismatches 122; Indels 158; Gaps 12;

QY 76 IYLLDDPHEALAAALIESAEHSLDLYQYIWRNDISGRLLFNLMLYLAERGVRVRLLD 135
Db 11 IQLENGEQYYPVAFKAIGEAQERIILETFIWFEDDVKGQJHAALLAAQGVKAEVLDD 70
QY 136 DNNTRGLDLDLALDHPNIEVLFNPFVLRKWRALGYLTDFFPLNRMRMINKSFTADNRA 195
Db 71 GYGSPDLSDFVNLTAAGVVFRIYDP-----RPLFGWRTN---VFRMRKIVVIDARI 123
QY 196 TILGGRNIGDEYFKVGEDTVFADLDILATGVSUVGEVSHDFD-----RYWASHSA 244
Db 124 AFIGGLNYSAEHMSYSGPEAKQDYAVRLEGPVIEDIL-QFELENLPQCSAARRWRRH-- 180
QY 245 HNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVOTRL 304
Db 181 HKAENRQPE---AQANGY-----AVRF----- 202
QY 305 ISDTPAKGLDRDRRKPPIAGRLQDALQKPEKSVYLVSPYFVPTKSGTDLAKLVQDGDIV 364
Db 203 -----GRL----- 205
QY 365 TVLTNSLOATDVAHVSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHA 424
Db 206 -----SGYY-----LVKGGVQVFEYRR-----PLHG 227
QY 425 KTFIVDGKRIFGSNLDRPSARLNTMGVYVIESPKIAEQMERTLAD 471
Db 228 KVALMDDHWATVGSSNLDPLSLNLEANVLIHD-----RHFNOTLRD 270

RESULT 34
US-10-450-763-31510
; Sequence 31510, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 31510
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (456)..(476)
; OTHER INFORMATION: Phospholipase D Active site proteins motifs. domain
; OTHER INFORMATION: identified by eMATRIX, accession number PF00614B, p-value=4.774e-
; OTHER INFORMATION: 14, raw score of 14.45
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (613)..(714)
; OTHER INFORMATION: Sodium:sulfate symporter transmembran domain identified by
; OTHER INFORMATION: Pfam, accession name Na_sulph_sym, E-value=6.1e-46, Pfam score c
; OTHER INFORMATION: 163.9
US-10-450-763-31510

Query Match 6.4%; Score 175; DB 5; Length 1097;
Best Local Similarity 20.4%; Pred. No. 7.1e-06;
Matches 83; Conservative 44; Mismatches 122; Indels 158; Gaps 12;

QY 76 IYLLDDPHEALAAALIESAEHSLDLYQYIWRNDISGRLLFNLMLYLAERGVRVRLLD 135
```

```
Db 241 IQLENGEQYYPVAFKAIGEAQERIILETFIWFEDDVKGQJHAALLAAQGVKAEVLDD 300
QY 136 DNNTRGLDLDLALDHPNIEVLFNPFVLRKWRALGYLTDFFPLNRMRMINKSFTADNRA 195
Db 301 GYGSPDLSDFVNLTAAGVVFRIYDP-----RPLFGWRTN---VFRMRKIVVIDARI 353
QY 196 TILGGRNIGDEYFKVGEDTVFADLDILATGVSUVGEVSHDFD-----RYWASHSA 244
Db 354 AFIGGLNYSAEHMSYSGPEAKQDYAVRLEGPVIEDIL-QFELENLPQCSAARRWRRH-- 410
QY 245 HNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVOTRL 304
Db 411 HKAENRQPE---AQANGY-----AVRF----- 432
QY 305 ISDTPAKGLDRDRRKPPIAGRLQDALQKPEKSVYLVSPYFVPTKSGTDLAKLVQDGDIV 364
Db 433 -----GRL----- 435
QY 365 TVLTNSLOATDVAHVSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHA 424
Db 436 -----SGYY-----LVKGGVQVFEYRR-----PLHG 457
QY 425 KTFIVDGKRIFGSNLDRPSARLNTMGVYVIESPKIAEQMERTLAD 471
Db 458 KVALMDDHWATVGSSNLDPLSLNLEANVLIHD-----RHFNOTLRD 500

RESULT 35
US-10-474-792-354
; Sequence 354, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 354
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-354

Query Match 6.2%; Score 168.5; DB 5; Length 525;
Best Local Similarity 20.0%; Pred. No. 8.8e-06;
Matches 89; Conservative 63; Mismatches 131; Indels 163; Gaps 20;

QY 96 ABHSLDQYI-----WRNDISGRLLFNLMLYLAERGVRVRLLDNNTGRGLDDL-LLA 148
Db 185 AKYIEFLEFIIAEGQMW-----GEIL-SILEKKVSEGVEVRVLF-----GNLSTLS 233
QY 149 LD-----SHPNIEVRLF---NPFVLRKWRALGYLTDFFPLNRMRMINKSFTADNRAITLGG 200
Db 234 SDYAKELEIQIGIKAKSFLPISPFIITY-----NYDRHRKIVVIDGEVSPFTGG 281
QY 201 RNIGDEYF-KVGEDTVFADLDILATGVS-----VGEVSHDFDPRYWASHSAHN 246
Db 282 INLADEYINEVERFHWKAGLMEGEATDSFILFLQWMSITEKEILIDPYLSDHSLK- 340
QY 247 ATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVOTRLIS 306
Db 341 -----LPSDGY-----VIPY-----G 351
QY 307 DTPAKGLDRDRRKPPIAGR--LDALKQKPEKSVYLVSPYFVPTKSGTDLAKLVQDGDIV 364
Db 352 DSP---LDTDK-----IGKNYVIDILNHAKEYVYIMTPYLIILDSEMEHALRFASERGVDI 403
```

```
Qy 365 TVLNSLOATVAAVHSGYVYKRPDLKAGIKLYELQPNHAVPATKDKLGTGSSVTSLSHA 424
Db 404 RIIMPGV--PKGVPYALAKTYKALMSSGVKIYEQGF-----VHS 444
Qy 425 KTFIVDGKRIFGSNFIDPRARLNTENGVIIESPKIAEQMERTLADTTPYAYRVTL-- 482
Db 445 KVFISDNTKAVVGINLDYRSLYHHFECAT-----YLIRVSVIA 483
Qy 483 -----DKHNRLOHMDPATRKY 499
Db 484 DIVDNFEAQXQSLMTSDHLTQRPW 509

RESULT 36
US-10-312-273-75
; Sequence 75, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035W0
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 75
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-75

Query Match 5.2%; Score 141.5; DB 4; Length 476;
Best Local Similarity 22.8%; Pred No. 0.0019;
Matches 86; Conservative 51; Mismatches 137; Indels 103; Gaps 19;

Qy 145 LLLAL-DSHPNIEVLFNPFVLRKWRALGYLTDPPRLNR-----RMHNSFTADNRAT 196
Db 110 LUKALKERHPN--RFFVVF-----TGCPPSTSLAPNVIEHMKLSIIDGKYC 155
Qy 197 ILGRNI-----GDEFK-----VGEDTVFADLDILATGSGVVG-EVSHDFDR 237
Db 156 ILGNTNFEFMTGDEPVEKVDNPRLFVSGVRRPLAFRDQDILMRSTAFGLQUREEYHK 215
Qy 238 YWA--SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVES-----PLYOKI 290
Db 216 QFAMWDYIAHH-----NWFIDNPEQFAGACPLTLEQAEETVFGPK- 258
Qy 291 QTGRIDMQSVQTRLISDTPAKGLDRDRKPPFIAGLQDALQKPEKSVYLVSYPFVPTKSG 350
Db 259 HEDLVLDSSKIRIVLGGP-----HDQPNFVTQYELKLIQAGARSVKLAHMYFIPKDEL 313
Qy 351 TDALAKLVQD-GIDVTVLTNSLOATDVAHVSGY-----VKY-----RKPLLKAGIKL 397
Db 314 LNALVDVSHNHVLSLTITNGCHELS-PAITGPAWGNRIINYFALLYGKRYPLWKWKF-C 371
Qy 398 YELQPNHAVPATKDKLGTGSSV-----TSLHAKTFTVDGKRIFIGSNLDPRSLNT 450
```

```
Db 372 EKLPYPERV-----SIYEFALWETQLHKKCMIIDDEIFVIGSYNFGKKSADFY 420
Qy 451 EMGVVIESPKIAEQMER 467
Db 421 ESIWVIESPEVAAKANK 437

RESULT 37
US-10-137-129A-6
; Sequence 6, Application US/10137129A
; Publication No. US20030124108A1
; GENERAL INFORMATION:
; APPLICANT: Frohman, Michael
; APPLICANT: Morris, Andrew
; APPLICANT: Engebrecht, Joanne
; TITLE OF INVENTION: No. US20030124108A1
; FILE REFERENCE: Onyx2004-DIV2
; CURRENT APPLICATION NUMBER: US/10/137,129A
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 09/536,224
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-137-129A-6

Query Match 4.9%; Score 133.5; DB 4; Length 932;
Best Local Similarity 18.6%; Pred No. 0.028;
Matches 101; Conservative 80; Mismatches 165; Indels 197; Gaps 25;

Qy 98 HSLDQYYIWRNDISGRLLFNLMYLAABERGVRVRLDDNNTRGLDLDLLALDS----- 151
Db 377 HSDD-----WRLDI-----MLKRAEAGVRVSVILLFKE-----VELALGINSGYSKRT 419
Qy 152 ----HPNIEVLFNPFVLRKWRALGYLTDPPRLNRHNSKFTADNRATILGSGNIG--- 204
Db 420 LMLLHPNIKV-MRHPDLVTLW-----AHHEKLLVVDVVVAFVGLGDLAFGR 464
Qy 205 --DEYFK---VGEDTVFADLDILATGS---VVGVSVDH-----PDR 237
Db 465 WDDVQYRLTDLGDPSEPHVLOTPTLGSDDPAATPDLSHNQFFWLKGYDNLITKDWVQLDR 524
Qy 238 YWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETV-----OSPLYQKIQT 292
Db 525 PFEDFIDRETTPRMPWRDVGVVHGVAAAD-LARHFIQRMNFTTKTKARYKTPLYPYL-- 581
Qy 293 GRIDQSVQVTEFLISDTP-----AKGLDRDRRKPPIAGRLQDA-----LK 331
Db 582 ----LPKSTSTANLFPMPGPGOCATVQVLRSDRWS-----AGTLENSILNAYLHTIR 631
Qy 332 QPEKSVLVSPYFVPTKSGTDALAKLVODGID-----VTVLT----- 368
Db 632 ESQHFLXIENOFFISCSGDRVLANKVGEIVDRILKAHQEQCFRVVILLPLLPFGEGDI 691
Qy 369 ----NSLQATDVAHVSGY-----VKYRPLLKAGIKLYELQPNH 404
Db 692 STGGNSIQ-----AILHFTYRTLRCRGEHSILHRLKAAMGTAWRDYMSCTGLRTHGELGGH 747
Qy 405 AVPATKDKLGTGSSVTSLSHAKTFTVDGKRIFIGSNLDPRSL-ARLNTMGVWVIESPKIA 462
Db 748 PI-----SLEIYIHSKMLADDTVTIIGSANIDRSLGKRDELAILKD----- 793
Qy 463 EQMERTLADTTPYAYRVTLDKHNR-----LOWHDPATRKTYPNEPEAKLWK 509
Db 794 TEMFPLMDGVGYQAGRFALSRLRCRFSVILGANTWPDLDLRDPVCDDEF-----QLWQ 847
Qy 510 RIA 512
Db 848 ETA 850
```

; APPLICANT: Glenn, Gary M.					
; APPLICANT: Warnock, Dale E.					
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION					
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME					
; FILE REFERENCE: 660088.465					
; CURRENT APPLICATION NUMBER: US/10/408,765A					
; CURRENT FILING DATE: 2003-04-04					
; NUMBER OF SEQ ID NOS: 3077					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 2203					
; LENGTH: 548					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-10-408-765A-2203					
Query Match 4.6%; Score 124.5; DB 4; Length 548;					
Best Local Similarity 18.0%; Pred. No. 0.079;					
Matches 96; Conservative 76; Mismatches 145; Indels 215; Gaps 24					
Qy	124	AERGVRVRLLLDDNTRGLDLDLALDS-----HPNIVRLFNPFVLKRWALGY	173		
Db	6	AEEGVRVSILLFKE-----VELALGINSYSKRALMLLPNIKV-MRRHPDQVTLM----	54		
Qy	174	LTFEPPLNRMRHNSFTADNRATILGG-----RNIGD-----	205		
Db	55	-----AHHEKLLVVDQVAFLGLDLAYGRWDDLHYRLTDLGSDSSASAASQPPTPR	105		
Qy	206	-----EYFKVGED-----	226		
Db	106	PDSPATPDLSHNGFFWLCKDYSNLTIKDWQLDRPPDFIDRETTPRMPWRDVGVVVHGL	165		
Qy	227	VGEVSHDFRYWASHSAHNATRIIRSGNICKGLQALGYNDTSRHALLRYRETVESPL	286		
Db	166	PARDLARHFIRQM-----NFTKTKA-----KYKTPTYPLLPKSTSTANQLPF	209		
Qy	287	YQKIQTGRIDWQSVOITFLISDTFAKLDRDRRKPIAGRLQDA-----LKQEPSVY	338		
Db	210	--TLPGGC--TTQV-----LRSVDRWS-----AGTLENSILNAYLHTIRESQHPLY	253		
Qy	339	LVSPYFPVPTKSGTDALAKEYDGID-----VTVLT-----NS	370		
Db	254	IENQFFISCSDGRTLNVKGDEIVDRLLKAHQGCYRVVVLPLLPFGFEGDISTGGNS	313		
Qy	371	LQATDVAAVHSYG-----VKYRKPLLKAGIKLYELOPNHAVPATKD	411		
Db	314	IQ-----AILHFTVRTLCRGYSILHLKAAMGTAWRDYISICGLRTHGELGGHPV----	364		
Qy	412	KGLTGSSVTSLHAKTFIVDGKRFIGSFNLDPRS---ARLNTMGVWIE----SPKI---A	462		
Db	365	-----SELIYIHKSLVIAADDRVTIIGSANINDRSLLGKRDSELAVLIEDTETPSLMNGA	419		
Qy	463	EQMERTLADTTPYAYVRVTLDKNR--LQHWDPATKTYTYPNEPAKLWKRIA	512		
Db	420	EYQAGRFALSLRKHCFGVLGANTRPDLDURDPICDDFF-----QLWQDMA	465		
RESULT 40					
US-10-424-599-215690					
; Sequence 215690, Application US/10424599					
; Publication No. US20040031072A1					
; GENERAL INFORMATION:					
; APPLICANT: La Rosa Thomas J					
; APPLICANT: Kovalic David K					
; APPLICANT: Zhou Yihua					
; APPLICANT: Cao Yongwei					
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated					
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement					
; FILE REFERENCE: 38-21(53223)B					
; CURRENT APPLICATION NUMBER: US/10/424,599					
; CURRENT FILING DATE: 2003-04-28					
; NUMBER OF SEQ ID NOS: 285684					
; SEQ ID NO 215690					
; LENGTH: 553					

RESULT 40
US-10-424-599-215690
; Sequence 215690, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215690
; LENGTH: 553

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:25:45 ; Search time 28 Seconds
(without alignments)
852.710 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAAILSLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*
1: /SIDSS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /SIDSS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SIDSS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SIDSS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /SIDSS5/ptodata/2/pubpaa/RCT_NEW_PUB.pep.*
6: /SIDSS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
7: /SIDSS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
8: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
9: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
10: /SIDSS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
11: /SIDSS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
12: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2712	99.7	525	9	US-10-467-657-5462
2	294.5	10.8	504	11	US-11-045-004-2518
3	287	10.6	482	11	US-11-045-004-2698
4	241.5	8.9	503	11	US-11-079-463-7133
5	232	8.5	442	9	US-10-793-626-1000
6	229.5	8.4	429	11	US-11-079-463-7814
7	227.5	8.4	485	11	US-11-074-176-10
8	226.5	8.3	396	9	US-10-510-386-238
9	226.5	8.3	399	9	US-10-510-386-30
10	220	8.1	479	11	US-11-098-686-10838
11	108.5	4.0	1263	9	US-10-485-517-127
12	105.5	3.9	1234	9	US-10-467-657-4224
13	104.5	3.8	489	9	US-10-242-586-42
14	104.5	3.8	489	9	US-10-242-586-42
15	104.5	3.8	489	9	US-10-243-116-42
16	104.5	3.8	489	9	US-10-243-136-42
17	104.5	3.8	489	9	US-10-243-189-42
18	104.5	3.8	489	9	US-10-243-215-42
19	104.5	3.8	489	9	US-10-243-236-42
20	104.5	3.8	489	9	US-10-243-298-42
21	104.5	3.8	489	9	US-10-243-304-42

22	104.5	3.8	489	9	US-10-243-338-42	Sequence 42, Appl
23	104.5	3.8	489	9	US-10-243-345-42	Sequence 42, Appl
24	104.5	3.8	489	9	US-10-243-357-42	Sequence 42, Appl
25	104.5	3.8	489	9	US-10-245-083-42	Sequence 42, Appl
26	104.5	3.8	489	9	US-10-247-015-42	Sequence 42, Appl
27	104.5	3.8	489	11	US-11-157-996-2	Sequence 15, Appl
28	104.5	3.8	506	11	US-11-226-701-15	Sequence 15, Appl
29	104	3.8	328	11	US-11-087-099-7553	Sequence 7553, Ap
30	103	3.8	495	11	US-11-074-176-266	Sequence 266, App
31	101.5	3.7	1072	11	US-11-096-568A-27848	Sequence 27848, A
32	101.5	3.7	1181	11	US-11-096-568A-27847	Sequence 27847, A
33	101.5	3.7	1189	11	US-11-096-568A-27846	Sequence 27846, A
34	101	3.7	833	7	US-09-941-095-85	Sequence 85, Appl
35	101	3.7	833	11	US-11-198-746-85	Sequence 85, Appl
36	101	3.7	833	11	US-11-198-657-85	Sequence 85, Appl
37	101	3.7	833	11	US-11-198-657-85	Sequence 85, Appl
38	101	3.7	1057	11	US-11-096-568A-30809	Sequence 30809, A
39	101	3.7	1166	11	US-11-096-568A-30808	Sequence 30808, A
40	101	3.7	1174	11	US-11-096-568A-30807	Sequence 30807, A
41	101	3.7	1423	11	US-11-188-298-17435	Sequence 17435, A
42	100	3.7	331	11	US-11-045-004-674	Sequence 674, App
43	100	3.7	423	11	US-11-087-099-9468	Sequence 9468, App
44	100	3.7	1159	9	US-10-055-877-139	Sequence 139, App
45	100	3.7	1574	11	US-11-212-443-179	Sequence 179, App

ALIGNMENTS

RESULT 1
US-10-467-657-5462
; Sequence 5462, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5462
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5462

Query Match	99.7%	Score	2712	DB	9	Length	525
Best Local Similarity	99.6%	Pred. No.	5.7e-215				
Matches	523	Conservative	1	Mismatches	1	Indels	0
Gaps	0						
Qy	1	MHTDPKIQAMPSTISPMKTRSLISLCLLCCSSWLPLEERTESRHNTPKPVLLDN	60				
Db	1	MHTDPKIQAMPSTISPMKTRSLISLCLLCCSSWLPLEERTESRHNTPKPVLLDN	60				
Qy	61	ILQIRHTPHNGLSDIYLLDDPHEALAAARALTESAEHSLDLOYYIWRNDISGRLLFNLM	120				
Db	61	ILQIRHTPHNGLSDIYLLDDPHEAPAAARALIESAEHSLDLOYYIWRNDISGRLLFNLM	120				
Qy	121	YLAERGVRVRLDLLDDNNTRGLDLDLALDSDHNIEVRLFNFPVLRKRWALGYLTDFPRL	180				
Db	121	YLAERGVRVRLDLLDDNNTRGLDLDLALDSDHNIEVRLFNFPVLRKRWALGYLTDFPRL	180				
Qy	181	NRMHNKSFTADNRATILGGRNIGDGYFKVGEDTVFADLDILATGSGVGVSHDFDRIYA	240				
Db	181	NRMHNKSFTADNRATILGGRNIGDGYFKVGEDTVFADLDILATGSGVGVSHDFDRIYA	240				

```
Qy 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYBETVEQSPLYOKIOTGRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYBETVEQSPLYOKIOTGRIDWQSV 300
Qy 301 QTRLSIDTPAKGLDRDRKPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDLAKLVQD 360
Db 301 QTRLSIDSPAKGLDRDRKPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDLAKLVQD 360
Qy 361 GIDVTVLNSLQATDVAHVSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLQATDVAHVSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRIFFIGSFNLDPRSARLNTENGVVIESPKIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFFIGSFNLDPRSARLNTENGVVIESPKIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNPEPAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNPEPAKLWKRIAANKILSLPIEGLL 525
```

RESULT 2

```
US-11-045-004-2518
; Sequence 2518, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESES, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEI, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: CHARBIT, ALAIN
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENIO, ENCARNIA
; APPLICANT: PABLOS, BETRIZ DE
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
; APPLICANT: HAUF, JORG
; APPLICANT: ROSE, MATTHIAS
; APPLICANT: VOSS, HAMUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
```

```
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2518
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-11-045-004-2518

Query Match 10.8%; Score 294.5; DB 11; Length 504;
Best Local Similarity 23.5%; Pred. No. 4.1e-16;
Matches 119; Conservative 72; Mismatches 207; Indels 109; Gaps 17;

Qy 19 KTRSLISLCLLCSWLPPLERTESRHFNTSKPVLNDNIQIRHTPHNGLSDIYL 78
Db 96 KTK-LINAIHAI PNNTNEKLPRLSKRI--AHLTSIEPI-----KG-NKIEI 137
Qy 79 LDDPHEALAAARAAALIESAHSLDLQYIWRNDISGRLLFNLMYLAERGVVRLLDDNN 138
Db 138 LTNGEETFPVLLDALRKAENHIHQYIFKTDALSTEIRDILVEKAKSGVEVRFMDGLG 197
Qy 139 TRGLDDLLALLDHPNIEVRLFNPFVLKWRPALGYLTDFFPLNRMRMKNKSTADNRATIL 198
Db 198 SSKLGKAPLAPLKEAGVSIHAFDP-IASPMIV-----RTANLRNHRKIVIDGQIGFT 249
Qy 199 GGRNIGDEY-----FKVGEDTVPADLDILATGSGWGEVSHDFDRYWA--SHSAHNATRI 250
Db 250 GGLNIGEYRSNTPDFRVRWDT-----HIKITGAVIELQESFLNDWVYMENQAGAADGF 304
Qy 251 IRSNIGKGLQALGYNDTSRHALLRYRETVEQSPLYOKIOTGRIDWQSVQTRLSIDTPA 310
Db 305 ISESGSKQYFSPVDMGDEWA-----QVIYGGPYD----KEKWVRDS-- 341
Qy 311 KGLDRDRKPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDLAKLVQDGDIVTLTNS 370
Db 342 -----MLDLDSAKESVMIYSPYFVPTDEESLAVIRRVAMSGVDVRVITPG 386
Qy 371 LQATDVAHVSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVTSIHAKTFIVD 430
Db 387 --KDRGISFHGSNAYVKTMEAGAKMYAYADDSFV-----HAKAMLVLD 428
Qy 431 GKRIFIGSFNLDPRSARLNTENGVVIESPKIAEQMERTLADTTPEYAYRVTLDKHNRLQW 490
Db 429 GTRAAIGTANFDRSFRNLNHELMVFLYDE--SEAMHHLKRDFFKDF-----EDSRL-- 477
Qy 491 HDPATRKTYPNEPEAKLWKRIAANKILS 517
Db 478 ---FTMKDMENKPLLTTRIKEVLSSLLS 501
```

RESULT 3

```
US-11-045-004-2698
; Sequence 2698, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESES, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEI, WERNER
```


Db 157 RVQLLIKAQGVKRVVVYDAWMSGMATAWFDOLRKAGGK-----VLPFITSRNM 208
Qy 170 ALGYLTDFPRLNRRMHKSFTADNRATILGGRNIGDEYF---KVGEDTVFADLDILATG 225
Db 209 ITRY-----RINYHLHRKIVIDGKISWTGFGNIGDQVLGRKKFGH---WRDSQVRIVG 260
Qy 226 SVGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVQSP 285
Db 261 SASLLQERFVMDW-NASINNDDEIR-----FN-----ST 290
Qy 286 LYQKIQTGRIDWQSVOTRLISDTAKGLDRDRKPPPIAGRLQDALKOPEKSVYLVSFV 345
Db 291 LFPDLDEKNIHQDDVATQIISDGP-----DRYSYMWNGMRLMLLARNELWQTPYLI 344
Qy 346 PTKSGTDALAKLVQDGDVTVLTNSLOATDVAAVHSGVYKVRKPLKAGIKLYELQPNHA 405
Db 345 PDDAVFATWQFIAMSGVDVRIMIPC--KPDHPFIYRATQWANELTRFGVKIYIYEDGF- 401
Qy 406 VPATKDKGLTSSVTSLSHAKTFIVDGGKIFIGSNLDPRLSPRSARLNTMGVIE----SPKI 461
Db 402 -----LHAKTTIIDNFFSSVGMNQDYSYSLNFEDNAIFYDKNFKNKM 445
Qy 462 AEQMERTLADT---TPE 475
Db 446 AEAEEEDMKSHLLTPE 462

RESULT 8
US-10-510-386-238
; Sequence 238, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 238
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-238

Query Match 8.3%; Score 226.5; DB 9; Length 396;
Best Local Similarity 22.4%; Pred. No. 1.1e-10;
Matches 97; Conservative 58; Mismatches 177; Indels 101; Gaps 11;

Qy 74 SDIYLLDDPHEALAAARALIESAHSLSLDLOYIWRNDISGRLLFNMLAAERGVRVRL 133
Db 40 SDIELIHNGEDLCERLLDDIRQAESSVHVMFYIKNDDISLEFLKVLKDKAKSGVCVRL 99
Qy 134 LDDNTRGLDLDLALDHPNIEVRLFPVLRKRALGYLTDFPRLNRRMHKSFTADN 193
Db 100 IDRIGAMKVKKTLGSLKQSGVHVFFAN-----KPGPPFYRLNARNHRKIAVIDG 151
Qy 194 RATILGRNIGDEYF-KVGEDTVFADLDILATGVSVEVSH-----DFDRYWASHAHNAT 248
Db 152 KIGVGGFNIKEVGLGKAEFGPKDYLHRLMTGEGVADLQHIIFISDFK-----199
Qy 249 RIIRSGNIGKGLQALGYNDTSRHALLRYRETVQSP---LYOKIQTGRIDWQSVOTRLI 305
Db 200 -----REAPQAKPANSVFPFLOQGAVTHTHATKGF 230
Qy 306 SDTPAKGLDRDRKPPPIAGRLQDALKOPEKSVYLVSFVPTKSGTDALAKLVQDGDV 365
Db 231 S-----LEEKYISFIOAKERIMICTPYIIPSPALQOAVLSARERGIVS 275

Qy 366 VLTNSLOATDVAAVHSGVYKVRKPLKAGIKLYELQPNHAVPATKDKGLTSSVTSLSHAK 425
Db 276 VLVP--MKPDHPLVKEAAYTHFPALLKAGCYIYRY-----RGF-----YHAK 316
Qy 426 TFIVDGKRIFIGSNLDPRLSPRSARLNTMGVIESPK-----IAQMERTLADTTPPEYA 477
Db 317 ALIVDDRHRVMIGTSFNFRSLFLNDEVNVVHDKDWTQKQFPDVFVKESIEHAELLTKERYA 376
Qy 478 YRVTLDKHNRLOW 490
Db 377 KRPVMQR--PVEW 387

RESULT 9
US-10-510-386-30
; Sequence 30, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 30
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-30

Query Match 8.3%; Score 226.5; DB 9; Length 399;
Best Local Similarity 22.4%; Pred. No. 1.1e-10;
Matches 97; Conservative 58; Mismatches 177; Indels 101; Gaps 11;

Qy 74 SDIYLLDDPHEALAAARALIESAHSLSLDLOYIWRNDISGRLLFNMLAAERGVRVRL 133
Db 43 SDIELIHNGEDLCERLLDDIRQAESSVHVMFYIKNDDISLEFLKVLKDKAKSGVCVRL 102
Qy 134 LDDNTRGLDLDLALDHPNIEVRLFPVLRKRALGYLTDFPRLNRRMHKSFTADN 193
Db 103 IDRIGAMKVKKTLGSLKQSGVHVFFAN-----KPGPPFYRLNARNHRKIAVIDG 154
Qy 194 RATILGRNIGDEYF-KVGEDTVFADLDILATGVSVEVSH-----DFDRYWASHAHNAT 248
Db 155 KIGVGGFNIKEVGLGKAEFGPKDYLHRLMTGEGVADLQHIIFISDFK-----202
Qy 249 RIIRSGNIGKGLQALGYNDTSRHALLRYRETVQSP---LYOKIQTGRIDWQSVOTRLI 305
Db 203 -----REAPQAKPANSVFPFLOQGAVTHTHATKGF 233
Qy 306 SDTPAKGLDRDRKPPPIAGRLQDALKOPEKSVYLVSFVPTKSGTDALAKLVQDGDV 365
Db 234 S-----LEEKYISFIOAKERIMICTPYIIPSPALQOAVLSARERGIVS 278
Qy 366 VLTNSLOATDVAAVHSGVYKVRKPLKAGIKLYELQPNHAVPATKDKGLTSSVTSLSHAK 425
Db 279 VLVP--MKPDHPLVKEAAYTHFPALLKAGCYIYRY-----RGF-----YHAK 319
Qy 426 TFIVDGKRIFIGSNLDPRLSPRSARLNTMGVIESPK-----IAQMERTLADTTPPEYA 477
Db 320 ALIVDDRHRVMIGTSFNFRSLFLNDEVNVVHDKDWTQKQFPDVFVKESIEHAELLTKERYA 379
Qy 478 YRVTLDKHNRLOW 490
Db 380 KRPVMQR--PVEW 390

RESULT 10

```
US-11-098-686-10838
; Sequence 10838, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10838
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10838

Query Match      8.1%; Score 220; DB 11; Length 479;
Best Local Similarity 22.5%; Pred. No. 5,1e-10;
Matches 99; Conservative 59; Mismatches 160; Indels 122; Gaps 18;

Qy 72 GLSDIYLLDDPHEALAAARALIESAHSLSLDIQQYIWRNDISGRLLFNLMYLAERGVVR 131
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 GGNVYPLYNGEEVYPKMLEAISKATHRVYLSLTFIFGNDVFVGTFLFIKTINEAAIRGCEVK 179
Qy 132 LLLDDNTRGLDLLLLLADS-----HPNIEVRLP-NPEVLRKWRALGTLTDFPRL--NRR 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 ILLD-----GIGS-FFPIKSWREKLHPAIKLAYFLSPSLP-----PKLSINLR 222
Qy 184 MHNKSFTADNRATILGGRNIGDEYF-KVGEDTVFADLILATGSVVGVSHDFDRYWASH 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
223 NHRKILVCDSNIAFTGNNISQHLVTLTKSNRVQDIHFYCAGIAKOLEIAFLDWSFT 282
Qy 243 SAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVESQSPYQKIQOTGRIDWQSVOT 302
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
283 TG-----ETAR-----IPEDPIKKQ-----GSTLC 302
Qy 303 RLISDTPAKGLDRDRKPPPIAGRLDAL-----KOPEKSVLVSPYFVPTKSGTDALAKLV 358
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
303 RLIIIDGPGSPHE-----THDLFCAMSSSAQKNIRIMSPYFLFPTPOLSEALASAV 352
Qy 359 QDGDIDVTV---LTNSLQATDVAVAHVSQYKVRKPLLKAGIKLYELQPNHVAVPATKDKGLT 415
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
353 LRGVTVDIIIPLKNHHYLLQWMAHQ-----PQLVDKNHILFLOPPPPFA----- 397
Qy 416 GSSVTSLHAKTFIVDGKRIFTGSNLDPRSLNLTMGVIESPKIAEQMBERTLADTTPTE 475
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 -----HTKLLLDINDYTLMGSANLDPRSLNLP-----VVENFSTTLTNLIE 442
Qy 476 Y-----AYRVTLDDKHNRL 488
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
443 FFDKKKSQSYQLT--EHTKL 460
```

```
RESULT 11
US-10-485-517-127
; Sequence 127, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
```

```
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 02000349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 127
; LENGTH: 1263
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-127

Query Match      4.0%; Score 108.5; DB 9; Length 1263;
Best Local Similarity 19.8%; Pred. No. 3.4;
Matches 101; Conservative 64; Mismatches 143; Indels 203; Gaps 26;

Qy 42 EERTESRHFTSKPEVLNLIQIRHT-PHNNGLSDIYLLDDPHEALAAARALIESAHS- 99
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
776 QPKTAVKAYTVTKPQTQTVSKIAQVKPNNTGI-----RASVYEKTAKN 820
Qy 100 ---LDLQYIWRNDISGRLLFNLMYLAERGVVRLLDDNTRGLDLLLLLALSDHPNIE 156
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
821 AKYADRTFYVTKERAHGNETYVLL-----NNT-----SH-NIP 852
Qy 157 VRLFPFVLRKWRALG---VLTDFPRLNRRMHNKSF--ADNRATILGGRNIGDEYF-- 208
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
853 LGWENVVDLNV-QNLGKEVTKTKYTVNKSNNGLSMVPGTWKQVILTGNNAIOGTFNAT 911
Qy 209 ---KVGEDT-VFADL-----DILATGSVVGVSHDFDRYWASHSAHNATRIIRSG 254
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
912 KQSVGKDVVLYGTINNTGWVNAKDLTAPTAVKPTTSAKD-----YNYTVVIKNG 963
Qy 255 NIGKGLQALGYNDTSRHALLRYRE----TVEQSPL-----YQIKOTGRIDW----- 297
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
964 N---GYIYVTPNSDTAKYSLKAFNEQPFVAVVKEQVINGQTWYVYKLSNGKLAWIKSTDLA 1020
Qy 298 ---QSVQTRLISDTPAK---GLDRDRRKPPPIAGRLQDA----- 329
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1021 KELIKYNOTGMTLNOVAQIQAGLQYKPOQVQVPGKWTDAKENDVKHAMDTKRLAQDPALK 1080
Qy 330 ---LKQPE-----KSVYLVSPYFVPTKSGT 351
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1081 YQFLRLDQPNISIDKINQFLKKGVLNQGAAFNKAAQMYGINEVYLISHALLETGNGT 1140
Qy 352 DALAKLVQDGDIVT---VLTN-----LQATDVAVAHVSQYKVRK----- 388
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1141 SQLAK---GADVNNKVVNTSNKYHNVFGIAAYDNDPLREG-IKYAKQAGWDTVSKAI 1195
Qy 389 -----PLLKAGIK-LYELOPNHVPAT 409
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1196 VGGAKFTGNSVVKAGQNTLYKMRWNPAHPGT 1226

RESULT 12
US-10-467-657-4224
; Sequence 4224, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4224
; LENGTH: 1234
; TYPE: PRT
```

```
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4224

Query Match
Best Local Similarity 3.9%; Score 105.5; DB 9; Length 1234;
Matches 127; Conservative 63; Mismatches 190; Indels 235; Gaps 32;

Qy 39 PPLEERTESRHNTSK---PVLLDNIILQIRHTPHNGLSDIYLLDDPHEALAAALIES 95
Db 4 PIPKPREKSRWNLQSGLPLALARYL-----PHKRL--KAVLTQD 42
Qy 96 AEHSLDLO--YYIWRNDISGRLLFNLMYLAERG-----VVRVLLDDNNTRGLDLLL 147
Db 43 AEQALRLQTAWRFRPHDHTAVFLPDWETLPHYERFSPHQDLVSRLSALWQIKSGAADVLF 102
Qy 148 ----ALDSDHNIEVRLFNPPVLAK--WRALGYLTDFPRLNRM-----HNKSFTADNR 194
Db 103 VPVATAMQKLPVP-----PFLAGRTFWLKTGOTLDIGRUKTDLVDAGYHNHSHVVAAGE 156
Qy 195 ATILGGRNIGDEYFKVGEDTVFA-----DLDLATGVSUVGEV-----SHDFD 236
Db 157 FAVRGG--IVD-LFPMGSETPYRIDLFDDEIDSIKTFDTQRTISPVSEIRLLPAHEFP 213
Qy 237 RYASHAHNAATRIIRSGNCKGKGLQALGYNDETSRHALLRYRETVEOSP-----LYQKIQT 292
Db 214 -----TDSEAQKIFRS-----RFREEVDCGNPDAAVYKAVSN 245
Qy 293 GRID-----WQSVQTRLSIDTPAK 311
Db 246 GHFGAGVEYVYLPFFENELETLFYIGEDALFVSLGVDHAEANRFWMDVKSRY-----AMAQ 302
Qy 312 GLDRDRKRPI-----AGRLQD-ALKOPEKSVYLVV--PYFVPVKSGTDLAKL 357
Db 303 G---DETYPPLLPQHLVYADVFAGRLKNYQVLPDVSGKAHSLPDLAVNRQSDDDPQAL 359
Qy 358 --VQDGDIVTVL-----TNSLOATVAAVHSGVYKVRKPL-----LK 392
Db 360 KDFQTAFDGRILLCAESLGRRETMLGFLQONGLKAKEPVS-DWGFLSAHEPLMITVAPLA 418
Qy 393 AGIKLYELQPN---HAYPATKDKGLTGSSTVSLHAKTFIVDGKREIFIGSFNLPDRSARLN 449
Db 419 YGFKLGLQSSQQOQTVFASGEQKAVTDQTERFASA-----TNPLPSPL- 463
Qy 450 TEMGVVIESPKIAEQMERTLAD-----TTPYAYRVTLDKHNLQWHPDPAKTKTYPNE 502
Db 464 -----PQEREQSAAVSDGLKAAAVSTESSLYLVASDLHQ-----TRQQSAPS 507
Qy 503 PEAKLW-----KRIAAK 514
Db 508 PVGEGWGEKAVAAQ 522
```

```
RESULT 13
US-10-242-586-42
; Sequence 42, Application US/10242586
; Publication No. US20060073548A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C21
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US/10/242,586
```

```
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-242-586-42

Query Match
Best Local Similarity 3.8%; Score 104.5; DB 9; Length 489;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPIYQKIQT-GRIDWQSVQTRLSIDTPAKGLDRDRKP-----PIAGRLQDALKOPEKSVYL 339
Db 60 SPAWEFLEAEARQDRSCQVLVLESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPYFYVPT-----KSGTDLAKLVQ---DGIDVTVLNS---LQATD--VAAV 379
Db 119 ASYVWSLTGPDIGVNDSSSQLGALLQKLOQLLGRNISLAVATSSPTLARTSTDLQVLA 178
Qy 380 HSGYVYKVRKPL--LKAGIKLYELQPNHVAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSPFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
Db 212 SANMDWRSLTQVKELGAVIYNCNSHLAQDLEKT 243
```

```
RESULT 14
US-10-242-902-42
; Sequence 42, Application US/10242902
; Publication No. US20060073549A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C54
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
```



```
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-136-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLSIDTTPAKGLDRDRRKP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPEAEARQORDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLDSTAQESVHV 118

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTN-----LQATD--VAAV 379
Db 119 ASYVWSLTGPDIGVNDSSSQLGALLQKQLLGRNLSLAVATSSPTLARTSTDLQVLAA 178

QY 380 HSGVVKRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSGFWVVDGRHIYMG 211

QY 438 SFNLDPRSARLNTMGVVI-ESPKIARQEMT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 17
US-10-243-189-42
; Sequence 42, Application US/10243189
; Publication No. US20060074033A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C44
; CURRENT APPLICATION NUMBER: US/10/243,189
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
```

```
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-189-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLSIDTTPAKGLDRDRRKP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPEAEARQORDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLDSTAQESVHV 118

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTN-----LQATD--VAAV 379
Db 119 ASYVWSLTGPDIGVNDSSSQLGALLQKQLLGRNLSLAVATSSPTLARTSTDLQVLAA 178

QY 380 HSGVVKRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSGFWVVDGRHIYMG 211

QY 438 SFNLDPRSARLNTMGVVI-ESPKIARQEMT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 18
US-10-243-215-42
; Sequence 42, Application US/10243215
; Publication No. US20060073551A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C27
; CURRENT APPLICATION NUMBER: US/10/243,215
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
```

```
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-215-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLISDTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPFVFPVT-----KSGTDALAKLVQ---DGIDVTVLNLS-----LOATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSQLGEALQKQLLGRNLSLATVSSPTLARTSTDLQVLAA 178
Qy 380 HSGVYKVRKPL--LKAGIKLYELOPNHAPVATKDKGLTGSSVTSLSHAKTFTVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243

RESULT 19
US-10-243-236-42
; Sequence 42, Application US/10243236
; Publication No. US20060073552A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C35
; CURRENT APPLICATION NUMBER: US/10/243,236
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
```

```
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-236-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLISDTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPFVFPVT-----KSGTDALAKLVQ---DGIDVTVLNLS-----LOATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSQLGEALQKQLLGRNLSLATVSSPTLARTSTDLQVLAA 178
Qy 380 HSGVYKVRKPL--LKAGIKLYELOPNHAPVATKDKGLTGSSVTSLSHAKTFTVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243

RESULT 20
US-10-243-298-42
; Sequence 42, Application US/10243298
; Publication No. US20060073553A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C49
; CURRENT APPLICATION NUMBER: US/10/243,298
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
```



```
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-298-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLIISDTPAKGLDRDRKRP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQORDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLNS-----LQATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSSQLEALLQKQLQLLGRNISLAVATSSPTLARTSTDLQVLAA 178
Qy 380 HSGVVKYRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRUTRGV-----LHSKFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMBERT 468
Db 212 SANMDWRSLSQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 21
US-10-243-304-42
; Sequence 42, Application US/10243304
; Publication No. US20060073554A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C42
; CURRENT APPLICATION NUMBER: US/10/243,304
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/157942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
```

```
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-304-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLIISDTPAKGLDRDRKRP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQORDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLNS-----LQATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSSQLEALLQKQLQLLGRNISLAVATSSPTLARTSTDLQVLAA 178
Qy 380 HSGVVKYRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRUTRGV-----LHSKFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMBERT 468
Db 212 SANMDWRSLSQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 22
US-10-243-338-42
; Sequence 42, Application US/10243338
; Publication No. US20060073579A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C52
; CURRENT APPLICATION NUMBER: US/10/243,338
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
```

```
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-338-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPYKQKIQT-GRIDMQSVQTRLSIDTPAKGLDRDRKP-PIAGRLQDALKQPEKSVYL 339
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 60 SPAWEPLAEARQORDSCQLVLESIP-QDLPSAAGSQAQLGQAWLQLLDTAQESVHV 118
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTVNS----LOATD--VAAV 379
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 119 ASYWSLTGPDIGVNDSSQLGEALLQQLLGRNISLAVATSSPTLARTSTDLQVLAA 178
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 380 HSGVYKVRKPL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSIHKATFIVDGKRIFIG 437
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEFWVVDGRHIYMG 211
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 23
US-10-243-345-42
; Sequence 42, Application US/10243345
; Publication No. US20060073555A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Wood, William
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C33
; CURRENT APPLICATION NUMBER: US/10/243,345
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-345-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPYKQKIQT-GRIDMQSVQTRLSIDTPAKGLDRDRKP-PIAGRLQDALKQPEKSVYL 339
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 60 SPAWEPLAEARQORDSCQLVLESIP-QDLPSAAGSQAQLGQAWLQLLDTAQESVHV 118
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTVNS----LOATD--VAAV 379
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 119 ASYWSLTGPDIGVNDSSQLGEALLQQLLGRNISLAVATSSPTLARTSTDLQVLAA 178
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 380 HSGVYKVRKPL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSIHKATFIVDGKRIFIG 437
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEFWVVDGRHIYMG 211
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 24
US-10-243-357-42
; Sequence 42, Application US/10243357
; Publication No. US20060073556A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Wood, William
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C55
; CURRENT APPLICATION NUMBER: US/10/243,357
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
```



```
; ORGANISM: Homo Sapien
US-10-247-015-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLISTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 60 SPAWPLEAEARQQRDSQQLVLESIP-QDLPSAAGSPSAQPLGQAWLQLDLDTAQS VHV 118

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTN-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 119 ASYWSLTGPDIGVNDSSQSGEALLQKQLLGRNISLAVATSPTLARTSTDLOVLAA 178

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPAFKDKGLTSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEKFWVDGRHIYMG 211

QY 438 SFNLDPRSARLNTMGVVI--ESPKIAEQMERT 468
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAODLEKT 243

RESULT 27
US-11-157-996-2
; Sequence 2, Application US/11157996
; Publication No. US20060003370A1
; GENERAL INFORMATION:
; APPLICANT: BAKER, KEVIN P.
; APPLICANT: CHEN, JIAN
; APPLICANT: DESNOYERS, LUC
; APPLICANT: GODDARD, AUDREY
; APPLICANT: GODOWSKI, PAUL J.
; APPLICANT: GURNEY, AUSTIN L.
; APPLICANT: PAN, JAMES
; APPLICANT: SMITH, VICTORIA
; APPLICANT: WATANABE, COLIN K.
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES
; FILE REFERENCE: 39780-3630R1C25C1
; CURRENT APPLICATION NUMBER: US/11/157,996
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 10/242,505
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/197,942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/US01/27099
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/140,653
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-157-996-2

Query Match          3.8%; Score 104.5; DB 11; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLISTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 60 SPAWPLEAEARQQRDSQQLVLESIP-QDLPSAAGSPSAQPLGQAWLQLDLDTAQS VHV 118

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTN-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :

; ORGANISM: Homo Sapien
US-10-247-015-42

Query Match          3.8%; Score 104.5; DB 11; Length 506;
Best Local Similarity 25.0%; Pred. No. 1.8;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLISTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 77 SPAWPLEAEARQQRDSQQLVLESIP-QDLPSAAGSPSAQPLGQAWLQLDLDTAQS VHV 135

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTN-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 136 ASYWSLTGPDIGVNDSSQSGEALLQKQLLGRNISLAVATSPTLARTSTDLOVLAA 195

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPAFKDKGLTSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 119 ASYWSLTGPDIGVNDSSQSGEALLQKQLLGRNISLAVATSPTLARTSTDLOVLAA 178

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPAFKDKGLTSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEKFWVDGRHIYMG 211

QY 438 SFNLDPRSARLNTMGVVI--ESPKIAEQMERT 468
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAODLEKT 243

RESULT 28
US-11-226-701-15
; Sequence 15, Application US/11226701
; Publication No. US20060009632A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Robison, Keith E.
; APPLICANT: White, David
; APPLICANT: Williamson, Mark W.
; APPLICANT: Cook, William James
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,
; FILE REFERENCE: MPI03-0210MNIM
; CURRENT APPLICATION NUMBER: US/11/226,701
; CURRENT FILING DATE: 2005-09-14
; PRIOR APPLICATION NUMBER: US/10/386,414
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 09/426,282
; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 09/668,266
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/330,970
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: 09/724,599
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/860,193
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/571,689
; PRIOR FILING DATE: 2000-05-16
; PRIOR APPLICATION NUMBER: 10/283,023
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/010,943
; PRIOR FILING DATE: 2001-12-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-226-701-15

Query Match          3.8%; Score 104.5; DB 11; Length 506;
Best Local Similarity 25.0%; Pred. No. 1.8;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLISTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 77 SPAWPLEAEARQQRDSQQLVLESIP-QDLPSAAGSPSAQPLGQAWLQLDLDTAQS VHV 135

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTN-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 136 ASYWSLTGPDIGVNDSSQSGEALLQKQLLGRNISLAVATSPTLARTSTDLOVLAA 195

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPAFKDKGLTSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 119 ASYWSLTGPDIGVNDSSQSGEALLQKQLLGRNISLAVATSPTLARTSTDLOVLAA 178

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPAFKDKGLTSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEKFWVDGRHIYMG 211

QY 438 SFNLDPRSARLNTMGVVI--ESPKIAEQMERT 468
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAODLEKT 243
```


QY 311 KGLDRDRKPPPIAGRLQDALKQPEKSVLVSPYFVP---TKSGTDALAKLVQDGDIVTVL 367
Db 555 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLGQGVPCIEKLSQAGVKIWL 607
QY 368 TNSLQATDVAHVSGYVKRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHAFTF 427
Db 608 TGDKTET---AINIGYA---CSLLREGMK-----QILVTL-----SSDIEALEKQG- 648
QY 428 IVDGKRIFIGSFN-----LDRPSARLNTM-GVVIESPKEAQWERTLAD 471
Db 649 --DKEAVAKASFQSIKKQLREGMSQTAAVTDSAKENSEMFLGVIDGKSLTYALDSKLEK 706
QY 472 TTPEYAYR 479
Db 707 EFLELAIR 714

RESULT 32
US-11-096-568A-27847
; Sequence 27847, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27847
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1181)
; OTHER INFORMATION: Ceres Seq. ID no. 2049239
US-11-096-568A-27847

Query Match 3.7%; Score 101.5; DB 11; Length 1181;
Best Local Similarity 21.2%; Pred. No. 11;
Matches 78; Conservative 61; Mismatches 154; Indels 75; Gaps 18;

QY 151 SHPNIEVRLFPFVLKRWALGYL-TDFPRLNRRMHNKSFTA---DNRATILGGRNIGDE 206
Db 492 NQPNAL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDAAAFVIASRELGE 545
QY 207 YFKVGEDTV-FADLDILATGVSVGESHDFRYWASHSAHNATRIIRSGN-----IGKGL 260
Db 546 FFSRSQTSISLHEIDHM-TGEKVDVRYELLHVLFFSSSRKMSVIVRNPNRLLLSKGA 604
QY 261 QALGYNDETSRHALLRYRETVEQSPLYOK--IQTGRIDMQSV-----QTRLIISDTPA 310
Db 605 DSVMPK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTYREIDEDEVIVWEEEFNNAKTL 663
QY 311 KGLDRDRKPPPIAGRLQDALKQPEKSVLVSPYFVP---TKSGTDALAKLVQDGDIVTVL 367
Db 664 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLGQGVPCIEKLSQAGVKIWL 716
QY 368 TNSLQATDVAHVSGYVKRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHAFTF 427
Db 717 TGDKTET---AINIGYA---CSLLREGMK-----QILVTL-----SSDIEALEKQG- 757
QY 428 IVDGKRIFIGSFN-----LDRPSARLNTM-GVVIESPKEAQWERTLAD 471
Db 758 --DKEAVAKASFQSIKKQLREGMSQTAAVTDSAKENSEMFLGVIDGKSLTYALDSKLEK 815
QY 472 TTPEYAYR 479
Db 816 EFLELAIR 823

RESULT 33

US-11-096-568A-27846
; Sequence 27846, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27846
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1189)
; OTHER INFORMATION: Ceres Seq. ID no. 2049238
US-11-096-568A-27846

Query Match 3.7%; Score 101.5; DB 11; Length 1189;
Best Local Similarity 21.2%; Pred. No. 12;
Matches 78; Conservative 61; Mismatches 154; Indels 75; Gaps 18;

QY 151 SHPNIEVRLFPFVLKRWALGYL-TDFPRLNRRMHNKSFTA---DNRATILGGRNIGDE 206
Db 500 NQPNAL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDAAAFVIASRELGE 553
QY 207 YFKVGEDTV-FADLDILATGVSVGESHDFRYWASHSAHNATRIIRSGN-----IGKGL 260
Db 554 FFSRSQTSISLHEIDHM-TGEKVDVRYELLHVLFFSSSRKMSVIVRNPNRLLLSKGA 612
QY 261 QALGYNDETSRHALLRYRETVEQSPLYOK--IQTGRIDMQSV-----QTRLIISDTPA 310
Db 613 DSVMPK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTYREIDEDEVIVWEEEFNNAKTL 671
QY 311 KGLDRDRKPPPIAGRLQDALKQPEKSVLVSPYFVP---TKSGTDALAKLVQDGDIVTVL 367
Db 672 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLGQGVPCIEKLSQAGVKIWL 724
QY 368 TNSLQATDVAHVSGYVKRPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHAFTF 427
Db 725 TGDKTET---AINIGYA---CSLLREGMK-----QILVTL-----SSDIEALEKQG- 765
QY 428 IVDGKRIFIGSFN-----LDRPSARLNTM-GVVIESPKEAQWERTLAD 471
Db 766 --DKEAVAKASFQSIKKQLREGMSQTAAVTDSAKENSEMFLGVIDGKSLTYALDSKLEK 823
QY 472 TTPEYAYR 479
Db 824 EFLELAIR 831

RESULT 34
US-09-941-095-85
; Sequence 85, Application US/09941095
; Publication No. US20060068383A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:

Db 599 FIAEGWLLVALDYQIELRVLAHLSGDNELIRVFQEGRDHETASWFMGVPREAVD-- 656
Qy 378 AVHSGYVYKRPPLKA-----GIKLYELQPNHAVP----- 407
Db 657 -----PLMRAAKTINFGVLYGMSAHLRSQELAIPIYEEAQAIFERYFQSPKVR 705
Qy 408 ----ATKDKGLTSSVTSLSHAKTFIVDGKRFIFGSLNLDPRSARLNTMGVVIESP---K 460
Db 706 AWIEKTLSEGRRGYVETLF-----GRRRYVPDLEARVKSREAAER-MAFNMPVRGT 757
Qy 461 IAEQWERTLADTTP---EYAVRVTLDKHNRLOQWHDPAIRKTYPNEPEAKLWKRIAANKLS 517
Db 758 AADLMKLVKLPFRLEEMGARMLLQVHDELVLPAKERA-----EAVARLAKEVMEGVYP 813
Qy 518 L-LPIE 522
Db 814 LAVPLE 819
RESULT 36
US-11-198-794-85
; Sequence 85, Application US/11198794
; Publication No. US20060035257A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: LYAMICHEV, VICTOR I.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 05-Aug-2005
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/11/198,794
; FILING DATE: 05-Aug-2005
; CLASSIFICATION:
; FILING DATE: 28-Aug-2001
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-11-198-794-85
Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4; Mismatches 188; Gaps 30;
Matches 118; Conservative 65

Qy 118 NLMYLAERGRV-----RLLDDNNTRGI---DLLLLADS----HPNIEVLFNPF 163
Db 321 DLLALAAAGGRVHRAPEPYKALRDLKEARGLLAKLSVLALREGLPPGDDPML----- 376

Qy 164 VLRKWRALGYLTD-----FPLNRRMHNKSF--ADNRATI-----LGGNIGDE----- 206
Db 377 -----LAVLLDSPNTTPEGVARRYGEWTEEAGEAALSERLAFANLWGLEGEERLLW 429
Qy 207 -YFKVGE--DTVPADLLDILATG-----SVVGEVS-----HDFDRYW 239
Db 430 LYREVERPLSAVLAHMB--ATGVRLDVAYLRALSLEVAGEIARLEAEVFRLAGHPFN--- 484
Qy 240 ASHSAHNATRIIRSGNIGKGLQALGYNDTSRHA-----LLRYRETV 281
Db 485 -LNSRDQLERVLFD---ELGLPAIGKTEKTKRSTSAVLEALREAHPIVEKILQYRELT 540
Qy 282 EQ-----SPLYQKI--OTGRIDWQSVOT-----RLISDTPAKGLDRDRKPKPIAGRIQDA 329
Db 541 KLKSTYIDPLPLDILHPRTRGLHTRFNQTATATGRLSSDP--NLQNIPIVPTPLQQRIRRA 598
Qy 330 LKQPEKSVYLVSPY-----FVPTKSGTDALAKLVQDGIDVTVLTS-----LQATDVA 377
Db 599 FIAEGWLLVALDYQIELRVLAHLSGDNELIRVFQEGRDHETASWFMGVPREAVD-- 656
Qy 378 AVHSGYVYKRPPLKA-----GIKLYELQPNHAVP----- 407
Db 657 -----PLMRAAKTINFGVLYGMSAHLRSQELAIPIYEEAQAIFERYFQSPKVR 705
Qy 408 ----ATKDKGLTSSVTSLSHAKTFIVDGKRFIFGSLNLDPRSARLNTMGVVIESP---K 460
Db 706 AWIEKTLSEGRRGYVETLF-----GRRRYVPDLEARVKSREAAER-MAFNMPVRGT 757
Qy 461 IAEQWERTLADTTP---EYAVRVTLDKHNRLOQWHDPAIRKTYPNEPEAKLWKRIAANKLS 517
Db 758 AADLMKLVKLPFRLEEMGARMLLQVHDELVLPAKERA-----EAVARLAKEVMEGVYP 813
Qy 518 L-LPIE 522
Db 814 LAVPLE 819
RESULT 37
US-11-198-657-85
; Sequence 85, Application US/11198657
; Publication No. US20060040299A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: LYAMICHEV, VICTOR I.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 05-Aug-2005
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/09/941,193
; FILING DATE: 28-Aug-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-11-198-657-85
Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4; Mismatches 188; Gaps 30;
Matches 118; Conservative 65


```
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-11-198-657-85

Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4;
Matches 118; Conservative 65; Mismatches 175; Indels 188; Gaps 30;

Qy 118 NMVLAERGVV-----RLLDNNTRGL---DDLALDS-----HPNIEVRLNPF 163
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 321 DLLAARAGRVRAPEPKALRDLKEARGLAKLDSVLAREGLGLPGDDPML----- 376
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 164 VLRKWRALGYLTD----FPRLNRRMHNKSF--ADNRATI-----LGGRNIGDE---- 206
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 377 -----LAYLDPSTNTPFEGVARRYGGEWTEEAGERAAALSERLFANLWGLEGEERLLW 429
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 207 -YFKVGE--DTVFADLDLATG-----SVVGEVS-----HDFRYW 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 430 LYREVERPLSAVLAME--ATGVRLDVAYLRALSLEVAGEIARLEAEVFRLAGHPFN--- 484
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 240 ASHSAHNATRIIRSGNICKGLQALGYNDETSRHA-----LRYRETV 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 485 -LNSRDQLERVLFD---ELGLPAIGKTEKTKRSTSAVLEALREAHPIVEKILQYREL 540
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 282 EQ-----SPLYQKI--QTGRIDWQSVT-----RLISDTPAKGLDRRRKPPPIAGRLQDA 329
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 541 KLKSTYIDPLDLPRLHPTGRGLHTRFNQTATATGLSSDP--NLQNPVPTPLGQRIRRA 598
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 330 LKQPEKSVYLVSPY-----FVPTKSGTDALAKLVQGDIDVTVLTNS-----LQATDVA 377
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 599 FIABEGWLLVALDYSQIELRVLAHLSGDNILRVFQEGRDITHTETASWFMGVPREAVD-- 656
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 378 AVHSGYVYKRPPLKA-----GKLYELOPNHNP----- 407
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 657 -----PLMRRAAKTNFVGLVGMGAHRUSQELAIPEBAQAFIERFQSPFKVR 705
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 408 ----ATKDKGLTGSVTSLSHAKTPIVDGKRIFIGSFNLDPRSRARLNTMGVWIESP--K 460
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 706 AWIEKTLLEGRRRGYVETLF-----GRRRYVPDLEARVKSVEAAER-WAFNWPVGRGT 757
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 461 IAEQMERTLADTTP---EYAYRVTLDKHNRLQWHDHPATRKTYNPEPEAKLWKRIAAIL 517
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 758 AADLMKLMVKLFPRLBEMGARMLLQVHDELVLPAKERA-----EAVARLAKEVMGVP 813
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 518 L-LPIE 522
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 814 LAVPLE 819
```

```
RESULT 38
US-11-096-568A-30809
; Sequence 30809, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30809
; LENGTH: 1057
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)-(1057)
; OTHER INFORMATION: Ceres Seq. ID no. 4973364
US-11-096-568A-30809

Query Match 3.7%; Score 101; DB 11; Length 1057;
Best Local Similarity 21.3%; Pred. No. 11;
Matches 76; Conservative 60; Mismatches 153; Indels 68; Gaps 18;

Qy 151 SHPNIEVRLNPFVLRKWRALGYL-TDFFRLNRRMHNKSF--ADNRATILGGRNIGDE 206
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 383 NQPNAL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDFAAFVIASRELGFE 436
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 207 YFKVGEDTV-FADLDILATGVSVVGEVSHDFRYWASHSAHNATRIIRSGN-----IGKGL 260
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 437 FFSRSQTSISLHEIDM---TVYELLHVLE---FSSSRKRMVIVRNPNRLLLSKGA 489
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 261 QALGYNDETSRHALRYRETVESPLYQK--IQTGRIDWQSV-----QTRLISDTPA 310
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 490 DSVNFK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTITYREIDEDVYVWEEFNAKTL 548
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 311 KGLDRRRKPPPIAGRLQDALKQPEKSVYLVSPYFVP---TKSGTDALAKLVQGDIDVTVL 367
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 549 VTEDRD-----ALIDAAADKIEKOLLILGSTAVEDKLGKVPDCIEKLSQAGVKIWWL 601
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 368 TNSLQATDVAHVSGYVYKRPPLKAGIK--LYELOPN--HAVPATKDKGLTGSVTSLSH 423
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602 TGGKET--AINIGVA---CSLLREGMKQILVTLDSDDIEALEKQGDKEAVAKLREGMS 655
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 424 AKTFIVDGKRIFIGSFNLDPRSRARLNTMGVWIESPKIAEQMERTLADTTPYAYR 479
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 656 QTAATVD-----NSAKENSEMFLVDGKSLTYALDLSKLEFLELAIR 699
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 39
US-11-096-568A-30808
; Sequence 30808, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30808
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1166)
; OTHER INFORMATION: Ceres Seq. ID no. 4973363
US-11-096-568A-30808

Query Match 3.7%; Score 101; DB 11; Length 1166;
Best Local Similarity 21.3%; Pred. No. 12;
Matches 76; Conservative 60; Mismatches 153; Indels 68; Gaps 18;

Qy 151 SHPNIEVRLNPFVLRKWRALGYL-TDFFRLNRRMHNKSF--ADNRATILGGRNIGDE 206
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 492 NQPNAL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDFAAFVIASRELGFE 545
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 207 YFKVGEDTV-FADLDILATGVSVVGEVSHDFRYWASHSAHNATRIIRSGN-----IGKGL 260
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 546 FFSRSQTSISLHEIDM---TVYELLHVLE---FSSSRKRMVIVRNPNRLLLSKGA 598
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 261 QALGYNDETSRHALRYRETVESPLYQK--IQTGRIDWQSV-----QTRLISDTPA 310
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 599 DSVNFK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTITYREIDEDVYVWEEFNAKTL 657
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 311 KGLDRRRKPPPIAGRLQDALKQPEKSVYLVSPYFVP---TKSGTDALAKLVQGDIDVTVL 367
```

Db	658	VTEDRD-----ALIDAADAKIEDLLLGSTAVEDKLOQGVDPDCIEKLQSAGVIMWL	710
Qy	368	TNSLQATDVAAVHSGVYKYRPELLKAGIK--LYELQPN--HAVPATKDGLTGSSVTSLSH	423
Db	711	TGDKTET---AINGYA---CSLLREGMKQILVTLDSDIDEALEKQGDKRAVAKLREGS	764
Qy	424	AKTFIVDGKRIFIGSFNLDPPRSARUNTEM-GWIESPKIAEQMERTLTADTTPEYAYR	479
Db	765	QTAAAVTD-----NSAKENSMFGLVTDGKSILTVALDSKLEKEFLFLAIR	808

```

RESULT 40
US-11-096-568A-30807
; Sequence 30807, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30807
; LENGTH: 1174
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1174)
; OTHER INFORMATION: Ceres Seq. ID no. 4973362
US-11-096-568A-30807

```

Query Match	3.7%;	Score 101;	DB 11;	Length 1174;
Best Local Similarity	21.3%;	Pred. No. 12;		
Matches	76;	Conservative 60;	Mismatches 153;	Indels 68; Gaps 18;
Qy	151	SHPNIEVRLFNPFVLRKWRALGYL-TDPPLNRMRHMKSFTA---DNRATILGCRNITGDE	206	
Db	500	NQPNAEI-----IQKFRVLAICHTAIPDNVNSTGEITYEAESPDAAFVIASREILGFE	553	
Qy	207	YFKVGEDTV-FADLDILATGSSVGEVSHDFDQRYWASHAHNATRIIRSGN-----IGKGL	260	
Db	554	FFSSQTSISLHEIDHM---TVYELHLVLE---FSSSRKMSVIVANPENRLLLSKGA	606	
Qy	261	QALGYNDETGRHALLRVRVEQSPQLYQK--IQTGRIDWQSV-----QTRLISDTPA	310	
Db	607	DSVMFK-RLAKHGRQNRNETKEHTKKYAEAGRLTVITRYIDEDEYIVWEEEFNAKTL	665	
Qy	311	KGLDRDRRKPIAQLQDALQKPEKSVYLVSPYFVP---TKSGGTDAKLAVODGIDVTVL	367	
Db	666	VTDRD-----ALIDAAADKIEKDILIGSTAVEDKLOKGVDPCEKUSQAGVKTIWVL	718	
Qy	368	TNSLQATDVAAVHSGVYKRYKPLLKAGIK--LYELQPN--HAVPATKDKGITGTSVSTSLH	423	
Db	719	TGDKTET---AINIGYA---CSLLREGMKQILVTLDSDSIDIEALEKQGDKEAVAKLRGMS	772	
Qy	424	AKTIVDGKRIFIGSNLDRPSARLNTQM-GVVTIESPKIAEQMERTIADTTTPPEYAYR	479	
Db	773	QTAAVTD-----NSAKENSEMFGLVIDGKSLTVYALDQSKLEKFEPLAIR	816	

Search completed: May 2, 2006, 06:28:54
Job time : 31 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:06:20 ; Search time 232 Seconds
(without alignments)
1596.562 Million cell updates/sec

Title: US-10-665-990A-14

Perfect score: 2720

Sequence: 1 MHTDPKIQAMPSETISPMKT.....KMKRIAATLILPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2720	100.0	525	2	Q6W764_NEIGO
2	2712	99.7	525	2	Q5F887_NEIG1
3	2643	97.2	508	2	Q9JTT1_NEIMA
4	2573	94.6	525	2	Q9JYU0_NEIME
5	1182	43.5	524	2	Q7VUJ6_BORPE
6	1181	43.4	492	2	Q7WPE3_BORBR
7	1104.5	40.6	493	2	Q8FIS6_ECOL6
8	1102.5	40.5	493	1	YMDC_ECOL1
9	1102.5	40.5	493	2	Q83RUS_SHIFL
10	1100.5	40.5	493	2	Q7AFAB_ECO57
11	1098	40.4	528	2	Q5QUL0_SALCH
12	1093.5	40.2	493	2	Q8X917_ECO57
13	1092.5	40.2	514	2	Q9CKM5_PASMU
14	1089.5	40.1	494	2	Q8ZQ28_SALTY
15	1089.5	40.1	495	2	Q8Z7M0_SALTI
16	1089.5	40.1	495	2	Q5PGV9_SALPA
17	940.5	34.6	519	2	Q6F8A6_AC1AD
18	881.5	32.4	516	2	Q4LR64_9BURK
19	881	32.4	505	2	Q87JB2_VIBPA
20	872.5	32.1	501	2	Q8D641_VIBVU
21	871.5	32.0	524	2	Q8D861_VIBVU
22	869.5	32.0	554	2	Q7MMB3_VIBVU
23	862.5	31.7	501	2	Q8D800_VIBVU
24	850.5	31.3	520	2	Q8PIM4_XANAC
25	849.5	31.2	530	2	Q6MQI7_EDEAM
26	848.5	31.2	570	2	Q4FTL4_9GAMM
27	846	31.1	464	2	Q5QUY8_IDILO
28	843	31.0	520	2	Q4UWU7_XANCP
29	843	31.0	520	2	Q8P7A5_XANCP
30	828	30.4	523	2	Q6G185_BAROU
31	825	30.3	523	2	Q6G4X0_BARHE

RESULT 1

ID	Q6W764_NEIGO	466	2	Q98N23_RHILO	Q98n23 rhizobium 1
AC	Q6W764; PRELIMINARY;	518	2	Q8UEX3_AGRTS	Q8uex3 agrobacteri
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	466	2	Q984B3_RHILO	Q984b3 rhizobium 1
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	520	2	Q6FBS3_AC1AD	Q6fbs3 acinetobact
DE	Phospholipase D (Fragment).	529	2	Q9HTP4_PSEAE	Q9htp4 pseudomonas
GN	Name-pld;	774	2	Q62E43_BURMA	Q62e43 burkholderi
OS	Neisseria gonorrhoeae.	773	2	Q88CA5_PSEPK	Q88ca5 pseudomonas
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;	772.5	2	Q4LJD9_9BURK	Q4ljd9 burkholderi
OC	Neisseriaceae; Neisseria.	540	2	Q83J33_BURPS	Q83j33 burkholderi
OX	NCBI_TaxID=485;	550	2	Q6L1I3_PHOPR	Q6l1i3 photobacter
RN	[1]	510	2	Q88BC2_PSESM	Q88bc2 pseudomonas
RP	NUCLEOTIDE SEQUENCE.	542	2	Q4ZZX0_PSESY	Q4zzx0 pseudomonas
RC	STRAN=1291;	751	2	Q5E6I6_VIBF1	Q5e6i6 vibrio fisc
RX	MEDLINE=22935225; PubMed=14573659;	750	2	Q92QO3_RHIME	Q92qo3 rhizobium m
RY	DOI=10.1128/IAI.71.11.6381-6391.2003;				
RA	Edwards J.L., Entz D.D., Apicella M.A.;				
RT	"Gonococcal phospholipase d modulates the expression and function of complement receptor 3 in primary cervical epithelial cells.";				
RL	Infect. Immun. 71:6381-6391(2003).				
DR	EMBL; AV307929; AAQ77232.1; -; Genomic DNA.				
DR	GO; GO:0003824; F:catalytic activity; IEA.				
DR	InterPro; IPR001736; PLD.				
DR	Pfam; PF00614; PLDC; 2.				
DR	SMART; SM00155; PLDC; 2.				
DR	PROSITE; PS50035; PLD; 2.				
FT	NON TER 525 525				
SQ	SEQUENCE 525 AA; 59224 MW; 4B626F5264E076D1 CRC64;				
Query Match 100.0%; Score 2720; DB 2; Length 525;					
Best Local Similarity 100.0%; Pred. No. 7.6e-184;					
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCCSSWLPPLERTSRHNTSKPVLLDN	60		
Db	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCCSSWLPPLERTSRHNTSKPVLLDN	60		
Qy	61	ILQIRHTPHNGLSDIYLLDDPHEALAAALIESAEHSLDQYIWRNDISGRLLFNL	120		
Db	61	ILQIRHTPHNGLSDIYLLDDPHEALAAALIESAEHSLDQYIWRNDISGRLLFNL	120		
Qy	121	YLAERGVRLDDNNTRGLDILLALDSDHNEVRLNPPVLRKWRALGYLTDPPRL	180		
Db	121	YLAERGVRLDDNNTRGLDILLALDSDHNEVRLNPPVLRKWRALGYLTDPPRL	180		
Qy	181	NRRMHNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVGVSHDFDRY	240		
Db	181	NRRMHNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVGVSHDFDRY	240		
Qy	241	SHSAHWATRIIRSGNTGKGLQALGYNDTSRHALRYRETVEQSPLYQKIQTCRIDMQ	300		
Db	241	SHSAHWATRIIRSGNTGKGLQALGYNDTSRHALRYRETVEQSPLYQKIQTCRIDMQ	300		

```
Db 241 SHSAHNATRIIRSGNIGKGLQALQVNDTSRHALLRYRETVEQSPLYQKIQOTGRIDWQSV 300
Qy 301 QTRLISDTPAKGLDRDRRKPPPIAGRLQDALQKPEKSVYLVSPFYFPTKSGTDLAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRRKPPPIAGRLQDALQKPEKSVYLVSPFYFPTKSGTDLAKLVQD 360
Qy 361 GIDVTVLNSLQATDVAHVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLQATDVAHVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525

RESULT 2
Q5F887 NEIG1
ID Q5F887 NEIG1 PRELIMINARY; PRT; 525 AA.
AC Q5F887
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DE Putative phospholipase D-family protein.
GN OrderedLocusNames=NG00902;
OS Neisseria gonorrhoeae (strain ATCC 700825 / FA 1090).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=242231;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Lewis L.A., Gillaspay A.F., McLaughlin R.E., Gipson M., Ducey T.F.,
RA Ownbey T., Hartman K., Nydick C., Carson M.B., Vaughn J., Thomson C.,
RA Song L., Lin S., Yuan X., Najjar F., Zhan M., Ren Q., Zhu H., Qi S.,
RA Kenton S.M., Lai H., White J.D., Clifton S., Roe B.A., Dyer D.W.;
RA "The complete genome sequence of Neisseria gonorrhoeae.";
RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE004969; AAW89600.1; -; Genomic_DNA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 525 AA; 59244 MW; 67C126AFCDAA0A360 CRC64;

Query Match 99.7%; Score 2712; DB 2; Length 525;
Best Local Similarity 99.6%; Pred. No. 2.8e-183;
Matches 523; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLLEERTESRHNTSKPVLADN 60
Db 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLLEERTESRHNTSKPVLADN 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDQYYIWRNDISGRLLFNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDQYYIWRNDISGRLLFNLM 120
Qy 121 YLAERGVVRVRLLLDDNNTRGLDLLLLALDSDHPNIEVRLFPFVLRKWRALGYLTDFFRL 180
Db 121 YLAERGVVRVRLLLDDNNTRGLDLLLLALDSDHPNIEVRLFPFVLRKWRALGYLTDFFRL 180
Qy 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLILATGSVVGESHDPDRYWA 240
Db 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLILATGSVVGESHDPDRYWA 240
Qy 241 SHSAHNATRIIRSGNIGKGLQALQVNDTSRHALLRYRETVEQSPLYQKIQOTGRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKGLQALQVNDTSRHALLRYRETVEQSPLYQKIQOTGRIDWQSV 300
Qy 301 QTRLISDTPAKGLDRDRRKPPPIAGRLQDALQKPEKSVYLVSPFYFPTKSGTDLAKLVQD 360
```

```
Db 301 QTRLISDTPAKGLDRDRRKPPPIAGRLQDALQKPEKSVYLVSPFYFPTKSGTDLAKLVQD 360
Qy 361 GIDVTVLNSLQATDVAHVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLQATDVAHVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525

RESULT 3
Q9JTT1 NEIMA
ID Q9JTT1 NEIMA PRELIMINARY; PRT; 525 AA.
AC Q9JTT1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=NMA1646;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL: AL162756; CAB84874.1; -; Genomic_DNA.
DR PIR: B81859; B81859.
DR GO: GO:0003824; F: catalytic activity; IEA.
DR GO: GO:0008152; P: metabolism; IEA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 525 AA; 59305 MW; 1CFB7AC5B82F1B02 CRC64;

Query Match 97.2%; Score 2643; DB 2; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.1e-178;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLLEERTESRHNTSKPVLADN 60
Db 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLLEERTESRHNTSKPVLADN 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDQYYIWRNDISGRLLFNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAEHSLDQYYIWRNDISGRLLFNLM 120
Qy 121 YLAERGVVRVRLLLDDNNTRGLDLLLLALDSDHPNIEVRLFPFVLRKWRALGYLTDFFRL 180
Db 121 YLAERGVVRVRLLLDDNNTRGLDLLLLALDSDHPNIEVRLFPFVLRKWRALGYLTDFFRL 180
Qy 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLILATGSVVGESHDPDRYWA 240
Db 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLILATGSVVGESHDPDRYWA 240
```

QY 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSHALLRYRETVEQSPLYQKIOTGRIDWQSV 300
 Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSHALLRYRETVEQSPLYQKIOTGRIDWQSV 300
 QY 301 QTRLISDTPAKGLDRBRKPIAGRLQALQKPEKSVVLYSPYFVPTKSGTDALAKLVQD 360
 Db 301 QTRLISDDPAPKGLDRBRKPIAGRLQALQKPEKSVVLYSPYFVPTKSGTDALAKLVQD 360
 QY 361 GIDVTLVNSLOQATDVAHVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
 Db 361 GIDVTLVNSLOQATDVAHVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
 QY 421 SLHAKTFIVDGKRFIFGFSNLDPRSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRV 480
 Db 421 SLHAKTFIVDGKRFIFGFSNLDPRSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRV 480
 QY 481 TLDKHNRLQWHDPAKTRKTYNPEPEAKLWKRIAAILSLPIEGLL 525
 Db 481 TLDKHNRLQWHDPAKTRKTYNPEPEAKLWKRIAAILSLPIESLL 525

RESULT 4

Q9JYU0_NEIMB
 ID Q9JYU0_NEIMB PRELIMINARY; PRT; 508 AA.
 AC Q9JYU0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cardiolipin synthetase family protein.
 GN OrderedLocusNames=NM1434;
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OC NCBI_TaxID=491;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
 RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
 RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
 RA Dodson R.J., Nelson W.C., Gwin M.L., DeBoy R.T., Peterson J.D.,
 RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
 RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
 RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
 RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Masioni V., Piazza M.,
 RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
 RA Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002098; AAP41795.1; -; Genomic_DNA.
 DR FTR; B81083; B81083.
 DR TIGR; NM1434; -.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001736; PLD.
 DR Pfam; PF00614; PLDC; 2.
 DR SMART; SM00155; PLDC; 2.
 DR PROSITE; PS50035; PLD; 2.
 KW Complete proteome.
 SQ SEQUENCE 508 AA; 57350 MW; A2DE6CAC47CA25D5 CRC64;

Query Match 94.68; Score 2573; DB 2; Length 508;
 Best Local Similarity 97.88; Pred. No. 1.8e-173;
 Matches 497; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 18 MKTRSLISLLCLLSCSSWLPPLTEERTSRHFNFTSKPVLDDNLIQRHTPHNGLSDIY 77
 Db 1 MKTRSLISLLCLLSCSSWLPPLTEERTSRHFNFTSKPVLDDNLIQRHTPHNGLSDIY 60

QY 78 LLDDPHALAAALIESAHSLDQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 137
 Db 61 LLNDPHEFAAARAALIESAHSLDQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 120

QY 138 NTRGLDLLALLDSDHPNIEVLEPNFVLRKWRALGYLTDFPRLNRMHMKSFADNRATI 197
 Db 121 NTRGLDLLALLDSDHPNIEVLEPNFVLRKWRALGYLTDFPRLNRMHMKSFADNRATI 180
 QY 198 LGGNIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGNIG 257
 Db 181 LGGNIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGDIG 240
 QY 258 KGLQALGYNDTSHALLRYRETVEQSPLYQKIOTGRIDWQSVQTRILISDTPAKGLDRDR 317
 Db 241 KGLQALGYNDTSHALLRYRETVEQSPLYQKIOTGCIQWQSVTRILISDTPAKGLDRDR 300
 QY 318 RKPIAGRLQALQKPEKSVVLYSPYFVPTKSGTDALAKLVQDGIQVTLVNSLOQATDVA 377
 Db 301 RKPIAGRLQALQKPEKSVVLYSPYFVPTKSGTDALAKLVQDGIQVTLVNSLOQATDVA 360
 QY 378 AVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFG 437
 Db 361 AVHSGYVKYRPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFG 420
 QY 438 SFNLDPRSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVTLDKHNRLQWHDPAKTRK 497
 Db 421 SFNLDPRSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVTLDRHNRQLQWHDPAKTRK 480
 QY 498 TYPNEPEAKLWKRIAAILSLPIEGLL 525
 Db 481 TYPNEPEAKLWKRIAAILSLPIEGLL 508

RESULT 5

Q7VUJ6_BORPE
 ID Q7VUJ6_BORPE PRELIMINARY; PRT; 524 AA.
 AC Q7VUJ6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative phospholipase D protein.
 GN OrderedLocusNames=BP3092;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=520;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640420; CAE43360.1; -; Genomic DNA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001736; PLD.
 DR Pfam; PF00614; PLDC; 2.
 DR SMART; SM00155; PLDC; 2.
 DR PROSITE; PS50035; PLD; 2.
 KW Complete proteome.
 SQ SEQUENCE 524 AA; 56590 MW; EDD1DE78F59B466 CRC64;

Query Match 43.5%; Score 1182; DB 2; Length 524;
 Best Local Similarity 48.6%; Pred. No. 5.8e-75;
 Matches 258; Conservative 77; Mismatches 152; Indels 44; Gaps 10;

```
Qy 19 KTRSLISL--LCLLLCSCSSWLPLEERTESRHFNTSKPVLLDNILOIRHTP----- 68
Db 14 RAALLAUAAGCAILLGACT--LPPVDRTAGHALDAD-----QARATPLGGQVGE 62
Qy 69 --HNNGLSDIYLLDDPHEALAAARALISBSHSLDLQYIWRNDISGRLLFNLMYLAER 126
Db 63 ADTHPGLSGPHVLGDAQDAFAARMLLAARATLTDVQYIWRNDWTGTLQLQALHAAER 122
Qy 127 GVRVRLDDNNRTGLDLLLLALDSHPNIEVRLNPFVLRKWRALGYITDPRNRRMHN 186
Db 123 GVRVRLDDNNRTGLDLLLLALDSHPNIEVRLNPFVLRKWRALGYITDPRNRRMHN 182
Qy 187 KSFTADNRATILGGRNIGDEYFKVGEDTVFADLLDILATGVSVVGVSHDFDRYWASHAHN 246
Db 183 KSFTVDNQATILGGRNIGDEYFGATGVLFADLDVLAVGVVDVSAEFDAYWASEGAMP 242
Qy 247 ATRIIRSGNIGKGLQALGYNDTSRSH-----ALLRYRETVQSPLYQKIOTGRIDWQSVQ 301
Db 243 AGPLLLP-----PGAQTLRALAERARIEQDPAAGDYSALRELFFIRELMAGRLPLQWAP 298
Qy 302 TRLISDTPAKGLDRDRRKPIAG----RLQDALQPEKSVLVSPFYVPTKSGTDLAKL 357
Db 239 ARWYSDDPKGLG----KAPPAQLLTQQLRNILGEPRTLDLVSPFYVPAEAGTQAFAL 354
Qy 358 VQDGIDVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHAVPATKGLTGS 417
Db 355 ARGGAQVRVLNVALEATDVAVVHSGYAKRRKALLQAGVRLYEMRRSYAGPKQGRFGS 414
Qy 418 SVTSLHAKTFFVDGKRIFIGSFNIDPRSANLNTGTVIEPKIAEQMERTLADTTPEYA 477
Db 415 SGSSLHAKTFGVDGERVIFGSPFNPRSANLNTGTVIEPKIAEQMERTLADTTPEYA 474
Qy 478 YRVTLDKHNRLQW---HDPATKTYPNPEAKLMKRIAAKILSLPIEGLL 525
Db 475 YEVRLLDDGSLYWLQORDGATVR-HDSEPGVSLWRRFVSLPLPLEPL 524

RESULT 6
Q7WP63 BORBR PRELIMINARY; PRT; 492 AA.
AC Q7WP63
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative phospholipase D protein.
GN OrderedLocusNames=BB0822;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640439; CAE31321.1; -; Genomic DNA.
DR GO; GO:0003824; F:metabolic activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
```

```
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PSS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 492 AA; 53309 MW; BC6BFE8AC34E726 CRC64;

Query Match 43.4%; Score 1181; DB 2; Length 492;
Best Local Similarity 49.6%; Pred. No. 6.3e-75;
Matches 253; Conservative 70; Mismatches 147; Indels 40; Gaps 8;

Qy 38 LPLEERTESRHFNTSKPVLLDNILOIRHTPHNNGLSDI-----YLLDDPHEALA 87
Db 1 MPPVPDRTAGHALDAS-----QARATPLGGQIGDLADAHFGLSGFHALGDAQDAPA 51
Qy 88 ARAALISBSHSLDLQYIWRNDISGRLLFNLMYLAERGVVRVRLDDNNRTGLDLLLL 147
Db 52 ARMLAARATRLTDVQYIWRNDWTGTLQLQALHAAERGVVRVRLDDNNRTGLDLLLL 111
Qy 148 ALDSHPNIEVRLNPFVLRKWRALGYITDPRNRRMHNKSFADNRATILGGRNIGDEY 207
Db 112 ALDAHPNAEVRLNPFVLRKWRALGYITDPRNRRMHNKSFADNRATILGGRNIGDEY 171
Qy 208 FKVGEDTVFADLLDILATGVSVVGVSHDFDRYWASHAHNATRIIRSGNIGKGLQALGYND 267
Db 172 FGATDGVLFADLDVLAVGVVDVSAEFDAYWASEGAMPAGPLLP-----PGAQTLRALA 227
Qy 268 ETSRH-----ALLRYRETVQSPLYQKIOTGRIDWQSVQTLISDTPAKGLDRDRRKPI 322
Db 228 ERAARIQDDPAAGDYSALRELFFIRELMAGRLPLQWAPARWSDDDPAKGLG----KAPP 283
Qy 323 AG-----RLQDALQPEKSVLVSPFYVPTKSGTDLAKLVQDGDIVTLTNSLOATDVAA 378
Db 284 AGLLTQQLRNILGEPRTLDLVSPFYVPTKSGTDLAKLVQDGDIVTLTNSLOATDVAA 343
Qy 379 VHSYGVYKRPKLLKAGIKLYELOPNHAVPATKGLTGSSTVSLHAKTFFVDGKRIFIGS 438
Db 344 VHSYAKRRKALLQAGVRLYEMRRSYAGPKQGRFGSSGSSSLHAKTFFVDGGERVIFGS 403
Qy 439 FNLDPRSANLNTGTVIEPKIAEQMERTLADTTPEYAVRVTLDKHNRLQW---HDPAT 495
Db 404 FNFDPRSANLNTGTVIEPKIAEQMERTLADTTPEYAVRVTLDKHNRLQW---HDPAT 463
Qy 496 RKTYPNPEAKLMKRIAAKILSLPIEGLL 525
Db 464 VR-HDSEPGVSLWRRFVSLPLPLEPL 492

RESULT 7
Q8FIS6 ECOL6 PRELIMINARY; PRT; 493 AA.
AC Q8FIS6
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein ymdC.
GN Name=yndC; OrderedLocusNames=c1310;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O6.H1 / CFT073 / ATCC 700928;
RX MEDLINE=22398234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli".
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016759; AAN7983.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
```



```
Db 390 AKTFSIDGKTVIGSFNDFPRSTLLNTMGFVIESETLAQLDKRFTQSYDAAWQLRLD 449
QY 484 KHNRLQWHDPAKTR--TYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 450 RWGRINWVDRHAKKEIVLKKEPATSFWKRVVMVRLASILPVEWLL 493

RESULT 11
ID Q57QL0_SALCH PRELIMINARY; PRT; 528 AA.
AC Q57QL0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Putative phospholipase.
GN Name=ymdC; OrderedLocusNames=SC1095;
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698 (2005).
DR EMBL; AE017220; AAK65001.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 528 AA; 60029 MW; D517EE93D71CDBB CRC64;

Query Match 40.4%; Score 1098; DB 2; Length 528;
Best Local Similarity 44.2%; Pred. No. 5.2e-69;
Matches 235; Conservative 88; Mismatches 175; Indels 34; Gaps 8;

QY 17 PMKTRS-----LISLCLLLCCSCSWLPPLEERTESRHNFTSKPVLDDNILQIRHT--- 67
Db 8 PLKSPSGRFOILLPVTLCLNRYTSLVMVKLPGFTQD--YLLSKATTLPKTRLERAVEP 65
QY 68 --PHNGLSDIYLLDDPHEALARAALIESAHSLLDQYIWRNDISGRLLFNLMYLAEE 125
Db 66 LCAHPGECGIALDNLDAFAARYLTETAARTLDVQYITWEDMSGRLLSVLSAAK 125
QY 126 RGVVRLLDDNNTRGLDDLLALDSDHPNIEVLFNPFVLRKWFALGYLTDFFRLNRMH 185
Db 126 RGVHVRLLDDNNTPGLDDTLRLDSDHPNIEVLFNPFVLRKWFALGYLTDFFRLNRMH 185
QY 186 NKSPADNRATILGGRNIGDEYFKVGEDTVFADLDLATGSGVVEGSHDFDRYWASHAH 245
Db 186 NKSYTAGVTVLGVGRNIGDAYFGAGEEPLFSDLDVMAIGFVNDVANDERWRCSSVS 245
QY 246 NATRIIRSGNIGKGLQAL-----GYNDETSRHALLRYRETVEQSPLYQKIQTGRID 296
Db 246 TLQOVLSE-----QELTQRIELPESWYNDIETR----RYLHKLETSSQFMADLDRGLP 296
QY 297 WQSVQVTRLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVLYVSPFVPTKSGTDALAK 356
Db 297 LIWAKTELLSDPSKSGKGAQRHSLLPQRLFDVNGSPETERIDIISAYFVPTAGVAQLLN 356
QY 357 LVQDGDIVTVLTNSLOATDVAHVSGVYKRPKLLKAGIKLYELOPNHA--VPATKDKGLT 415
Db 357 LVKRGVKIALTNSLANDVAVHAGYARWRKLLRYGVELYELKPTREHETAVHQRGLT 416
QY 416 GSSVTSLHAKTIFVDGKRIFIGSNLDPDRSARLNTMGVGVIESPKIAEQMERTLADTPE 475
Db 417 GNSGSSLHAKTFTSDGSKVFIGSLNDFPRSTLLNTMGFVIESETLATLHKHFTQSQRD 476
QY 476 YAYRVTLDKNRLQWHD--PATKTYTNEPEAKLWKRIAAKILSLPIEGLL 525
Db 477 AAWQLRLDRGRINWIDROQEEKVLKKEPATSFWRQVRLVRLAAILPVEWLL 528
```

```
RESULT 12
Q8X9I7_ECO57
ID Q8X9I7_ECO57 PRELIMINARY; PRT; 493 AA.
AC Q8X9I7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative synthase.
GN Name=ymdC; OrderedLocusNames=z1680;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074933; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
DR EMBL; AE005174; AAG55792.1; -; Genomic_DNA.
DR PIR; D85666; D85666.
DR PIR; H90806; H90806.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 493 AA; 55981 MW; DCC21D8AAE643457 CRC64;

Query Match 40.2%; Score 1093.5; DB 2; Length 493;
Best Local Similarity 47.2%; Pred. No. 9.8e-69;
Matches 219; Conservative 84; Mismatches 144; Indels 17; Gaps 5;

QY 72 GLSDIYLLDDPHEALARAALIESAHSLLDQYIWRNDISGRLLFNLMYLAABRGVVR 131
Db 37 GCGCLFPLEKSLDAFAARYLAEMSEHTLDVQYIWDQDMSGRLLSFALLAAKRGVVR 96
QY 132 LLLDDNNTRGLDALLALDSDHPNIEVLFNPFVLRKWRALGYLTDFFRLNRMHNSFTA 191
Db 97 LLLDDNNTPGLDDTLRLDSDHPNIEVLFNPFVLRKWRALGYLTDFFRLNRMHNSFTV 156
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDLATGSGVVEGSHDFDRYWASHAHNATRII 251
Db 157 DGVTVLVGGRNIGDAYFGAGEEPLFSDLDVMAIGFVNDVADDFAFYWYCKSVPLQOVL 216
QY 252 RSGNIGK-----LQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDMQSVOTRL 304
Db 217 ---DVPEGEMADRIELPASHWINDAMTH----RYLRKMESSFHNLVVDGFLPIWAKTRL 269
QY 305 ISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVLYVSPFVPTKSGTDALAKLVQGDIDV 364
Db 270 LSDDPAKGEKAKRHSLLPQRLFDVNGSPETERIDIISAYFVPTAGVAQLLRVMVKVKI 329
QY 365 TVLTNSLOATDVAHVSGVYKRPKLLKAGIKLYELOPNHAVPAT-KDKGLTGSSTVSLH 423
Db 330 AILTNSLAANDVAVHAGYARWRKLLRYGVELYELKPTREQSSTLHDRGITGNSGASLH 389
QY 424 AKTFFIVDGRKRIFGSNLDPDRSARLNTMGVGVIESPKIAEQMERTLADTTPYAYRVYR 483
Db 390 AKTFSIDGKTVIGSFNDFPRSTLLNTMGFVIESETLAQLDKRFTQSYDAAWQLRLD 449
QY 484 KHNRLQWHDPAKTR--TYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 450 RWGRINWVDRHAKKEIVLKKEPATSFWKRVVMVRLASILPVEWLL 493
```

```
RESULT 13
Q9CKM5_PASMU PRELIMINARY; PRT; 514 AA.
ID Q9CKM5;
AC Q9CKM5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PM1586.
GN OrderedLocusNames=PM1586;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006196; AAK03670.1; -; Genomic DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome; Hypothetical protein.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 514 AA; 57882 MW; C65B0EB65D80F216 CRC64;

Query Match 40.2%; Score 1092.5; DB 2; Length 514;
Best Local Similarity 46.5%; Pred. No. 1.2e-68;
Matches 238; Conservative 78; Mismatches 173; Indels 23; Gaps 8;

QY 23 LLSLCLLLCSCSSLPLEERTESRHTSKPVLDMILQIRH-----TPHNGLSDIY 77
DB 17 LFSVLIALISYQRL--PTSRPISTH-----PPLNPNGLLARHLLPQVAQHPN-LTGLY 68
QY 78 LLDPPHEALAAARALIESAEHSLLDQYYIWRNDISGRLLFNLMYLAERGVRRVRLDDN 137
DB 69 PLGDKDAFLARLALSEHAETLDLQYYIWHNDVSGHLLQSLYKAAVRGVKVRLLDDN 128
QY 138 NTRGLDLLLLALDSDHPNIEVRLFNFPVLRKWRALGYLDFPFLNRRMNKSFADNRATI 197
DB 129 NTKGMDTTLASINAHPIQIRLFPFQORQYRWLGFLSDFPFLNRRMNKSFADGVMSI 188
QY 198 LGGRNIGDEYFKVGEDTVFADLDILATGVSVGEVSHDFRYWASHSAHNATRIIRSGNIG 257
DB 189 LGGRNIGDEYFDVGVNGVLFADLDVAATGAVTHIQTDPRYWNSSPYLESIIIRDPI 248
QY 258 KGLQALGYNDETSRRHALLRYRETVEQSPLYQKIQTGRIDWQSVOTRLISDTPAKGLDRDR 317
DB 249 PNPPLALDDETEQ-----TYLKQLTELPFAKSLKAGTLAFTWAEALISDDPKKALGKSL 303
QY 318 KPPTAGRLQDALKQPEKSVLYSPYFPTKSGTDLAKLVQDGDVTVTLNSLOATDVA 377
DB 304 TQDSVLAIAPTMLNAKNLIIVSYFPTTHVGVDFLSRSIQGTQGVTSILNLSLEATDVS 363
QY 378 AVHSGVYKRPKLLKAGIKLYELQPNHAPAPKDKG--LTGSSVSLHAKTFIVDGKRF 435
DB 364 IVHSGYAKHRTLLQKQQLYELKLP-HATIQMESSGHLLKGASSASLHAKTFILNRYLF 422
QY 436 IGSFNLDPRLSARLNTMGVWIESPKIAEQMERTLADTTPEYAYRVTLDKHNRLOWHPAT 495
DB 423 VGSFNWDPRLSAMLNTMGLLIDSPELARLLSDGLQONQANYAFSVKLNEAQAALYWTQEN 482
QY 496 RK--TYPNEPEAKLWKRAAKILSLPIEGLL 525
DB 483 GKWIYENEPHTSWFKRFSVWCLSLWLVVEHLL 514

RESULT 14
Q8ZQ28_SALTY PRELIMINARY; PRT; 494 AA.
ID Q8ZQ28;
AC Q8ZQ28;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative phospholipase.
GN Name=yndC; OrderedLocusNames=STM1148;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2."
RL Nature 413:852-856(2001).
DR EMBL; AE008749; AAL20078.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 494 AA; 56217 MW; 571C01EFC4FCB288 CRC64;

Query Match 40.1%; Score 1089.5; DB 2; Length 494;
Best Local Similarity 47.4%; Pred. No. 1.9e-68;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

QY 72 GLSDIYLDPPHEALAAARALIESAEHSLLDQYYIWRNDISGRLLFNLMYLAERGVRR 131
DB 38 GEGCLALDNLDAFAARYRLTEMAARLDVQYYIWEDDMSGRLLFSVLLSAAKGVHVR 97
QY 132 LLLDDNTRGLDLLLLALDSDHPNIEVRLFNFPVLRKWRALGYLDFPFLNRRMNKSFSTA 191
DB 98 LLLDDNTPGLDDTLRLDSDHPNIEVRLFNFPFRTLRALGYLDFPFLNRRMNKSFSTA 157
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVGEVSHDFRYWASHSAHNATRII 251
DB 158 DGVTVLGGRNIGDAYFGAGEEPLFSLDDVMAIGVNDVANDPERYWRCSVSTLQOVL 217
QY 252 RSGNIGKGLQAL-----GYNDETSRRHALLRYRETVEQSPLYQKIQTGRIDWQSVOT 302
DB 218 SLSE-----QELTQRIELPESWYNDIIR-----RYLHKLETSSQFMADLDGRGLPLIWA 268
QY 303 RLISDTPAKGLDRDRPKPIAGRLQDALKQPEKSVLYSPYFPTKSGTDLAKLVQDGI 362
DB 269 RLLSDSPSKGKQKRSLLPQRLFDVNGSPTEIRIDIIISAYFVPTPAGVQLLNLRKGV 328
QY 363 DVTVLNLSQATDVAHVHSGVYKRPKLLKAGIKLYELQPNHA-VPAFKDKGLTGSSVTS 421
DB 329 KIAITNLSLAANDVAVHAGYARWRKLLRGVLYELKPTREHETAVHDLRGLTGNSSG 388
QY 422 LHAFTFIVDGKRIIFGSFNLDPRLSARLNTMGVWIESPKIAEQMERTLADTTPEYAYRV 481
DB 389 LHAFTFISDGKVFISGLNFDPRSTLLNTMGFVIESSETLATLIHKRFTQSDAAWQLR 448
QY 482 LDKGNRLQWHD--PATRKYTNEPEAKLWKRAAKILSLPIEGLL 525
DB 449 LDRWGRINWIDRQOEEKVLKKEPATREWRQVRLVRLAAILPVEWLL 494
```

```
RESULT 15
Q827M0_SALTI
ID Q827M0_SALTI PRELIMINARY; PRT; 495 AA.
AC Q827M0; Q7C9A7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein STY1185.
GN OrderedLocusNames=STY1185, t1772;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN 1
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN 2
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyonanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627269; CAD08272.1; -; Genomic DNA.
DR EMBL; AE016840; A069395.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 56348 MW; 4555257D8DB9 CRC64;

Query Match 40.1%; Score 1089.5; DB 2; Length 495;
Best Local Similarity 47.4%; Pred. No. 1.9e-68;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

QY 72 GLSDIYLLDDPHEALAAALIESAHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
DB 39 GEGCILALDNLDAFAARYLTETMAARTLDVQYIWEEDMSGRLLFVLSAARGVHVR 98

QY 132 LLLDDNNTRGLDLLLALDHPNIEVRLNPFVLRKWRALGYLTDFFLRNRRMHKSF 191
DB 99 LLLDDNNTPGLDRTLRLDHPNIEVRLNPFVLRKWRALGYLTDFFLRNRRMHKSF 158

QY 192 DNRATILGRRNIGDEYFKVGEDTVFADLDILATGVSVVGVSHDFDRYASHSAHNATRII 251
DB 159 DGVVTLVGGRRNGIDAYFGAGEEPLFSDLDVMAIGPVVNDVANDFERYWRCSSYSTLQQVL 218

QY 252 RSGNIGKGLQAL-----GYNDETSHALLRYRETVEQSPLYOKIOTGRIDWQSVQT 302
DB 219 SLSE-----QELTQRIELPSWVNDIEITR-----RYLHKLETQFMADLRGLTLPITWAKT 269

QY 303 RLISDTPAKGLDRDRRPPFIAGRLQDALKQPEKSVLVLPVFPVTPKSGTDLAKLVQDGI 362
```

```
RESULT 16
QSPGY9_SALPA
ID QSPGY9_SALPA PRELIMINARY; PRT; 495 AA.
AC QSPGY9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein ymdC.
GN Name=ymdC; OrderedLocusNames=SPA1703;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN 1
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Dalehaunt K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid."
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV77627.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 56310 MW; 8BAADE58874983A2 CRC64;

Query Match 40.1%; Score 1089.5; DB 2; Length 495;
Best Local Similarity 47.4%; Pred. No. 1.9e-68;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

QY 72 GLSDIYLLDDPHEALAAALIESAHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
DB 39 GEGCILALDNLDAFAARYLTETMAARTLDVQYIWEEDMSGRLLFVLSAARGVHVR 98

QY 132 LLLDDNNTRGLDLLLALDHPNIEVRLNPFVLRKWRALGYLTDFFLRNRRMHKSF 191
DB 99 LLLDDNNTPGLDRTLRLDHPNIEVRLNPFVLRKWRALGYLTDFFLRNRRMHKSF 158

QY 192 DNRATILGRRNIGDEYFKVGEDTVFADLDILATGVSVVGVSHDFDRYASHSAHNATRII 251
DB 159 DGVVTLVGGRRNGIDAYFGAGEEPLFSDLDVMAIGPVVNDVANDFERYWRCSSYSTLQQVL 218

QY 252 RSGNIGKGLQAL-----GYNDETSHALLRYRETVEQSPLYOKIOTGRIDWQSVQT 302
```

```
Db 219 SLSE-----QELTORIELPESWYNDEITR-----RYLHKLETSQFMAWLDLDCGTLPLIWAKT 269
Qy 303 RLISDTPAKGLDRDRRKPPIAGRILODALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGI 362
Db 270 RLLSDPSKGEKQAQRHSLLPQRLFDVWMSPTERIDIISAYFVPTPAGVAGQLLMLVRKV 329
Qy 363 DVTVLTSLQNTDYAAVHSGVYKVRKPLKAGIKLYELOPNHA--VPATKDKGLTGSSVTS 421
Db 330 KIAITLNSLAANDVAVHAGYARWRKLLRYGVELYELKPTRREHETAVHQRGLTNGSS 389
Qy 422 LHAKTFTVDGKRIFIGSNLDPRLARLNTENGWVIESPKIAEQMERTLADTPPYAYRV 481
Db 390 LHAKTFTSDGSKVIGSNLDPRLARLNTENGWVIESPTLATLHKRFTQSORDAAWQLR 449
Qy 482 LDKHNRLQWHD--PATRKTYNPEPEAKLWKRIAKIILSLDPIEGLL 525
Db 450 LDRWGRINWIDRQOEKVKLKEPATRPFWRQVRLVRLAAILPVEWLL 495

RESULT 17
Q6F8A6 ACIAD
ID Q6F8A6 ACIAD PRELIMINARY; PRT; 519 AA.
AC Q6F8A6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative phospholipase D protein.
GN OrderedLocusNames=ACIAD3001;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779 (2004).
DR EMBL; CR543861; CAG69709.1; -; Genomic_DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 519 AA; 59913 MW; C87537724E490292 CRC64;

Query Match 34.6%; Score 940.5; DB 2; Length 519;
Best Local Similarity 38.2%; Pred. No. 7.2e-58;
Matches 207; Conservative 101; Mismatches 177; Indels 57; Gaps 10;

Qy 18 MKTSLSLCLLCSCSSWLPPEERTESHFNTSKPVLNDILQIRHPTHNNGLSDIY 77
Db 1 MLTSLVTVGVINGCSTLPKHGP-EVTQYADIDTSQTSQAIIITPLKQ--QNLGLTGYH 57
Qy 78 LLDPHPEALAAALIESAHSLSLDLQYIWRNDISGRLLFNLMYLAERGVVRVRLLDN 137
Db 58 MLYEPLAALAAALRIDKAETLDLQYIWRNDIKIGALHALTRAADRGVVRVRLLDN 117
Qy 138 NTRGLDDLLALDGHNPTEVRLFPVLRKWRALGYLTDFPRLNRRMHNKSFTADNRATI 197
Db 118 NAKQIEGVLLALSQHONIEVKLFNPFYRFRKYRAMDMILDKRINRRMHNKSFIADNQVAL 177
Qy 198 LGGRNIGDEYFKVEDTVPADLDILATGTSVGEVSHDFDRFWASHAHNAITRISGNIG 257
Db 178 IGGRNMTQYNNVSDNYQFSDVLLVGAADVDDIVKSFDDYWNDEYAVSQNVIV----- 231
Qy 258 KGLQALGYNDTSRHALLRYRETVEQ-SPLYOKIQTG-----RID 296
```

```
Db 232 -----NFEQHR---LRYOSLKQQLDDYQEVTVQVNYLDLTTTRSHAFQWLNNNQLD 280
Qy 297 WQSVQTRLIISDTPAKGLDRDRRKPPIAGRILODALKQPEKSVYLVSPYFVPTKSGTDALAK 356
Db 281 W--VKATVVKDAPDKIRAKANKEQHNLNFQLIKHEKPEDNDVLISAYFVPEKQAKHLS 338
Qy 357 LVQDGIDVTVLTNSLQNTDYAAVHSGVYKVRKPLKAGIKLYELOP-----NHA 405
Db 339 MAQDGIKRVLTNSFKANDPVVHAFYAKYRQELLENGVOLYBFLPTLEKEELDXYTDEL 398
Qy 406 VPATKDKGLTGSSVTSLHAKTFTVDGKRIFIGSNLDPRLARLNTENGWVIESPKIAEQM 465
Db 399 VKAK-VSIKGLSRSSLHAKWMALDQKQVIGSFNFPDRSAYLNTLGVILNSPLAMAV 457
Qy 466 ERTLADTPPYAYRVLTLDKHNRLQWHDPA--RTKTYNPEPEAKLWKRIAKIILSLDPIEG 523
Db 458 HQTMDQQLTHYAYKLVLDANHKINWYRQTSTGTKIYTKPRMKWQKAGIKLISWLPTEG 517
Qy 524 LL 525
Db 518 FM 519

RESULT 18
Q4LR64_9BURK
ID Q4LR64_9BURK PRELIMINARY; PRT; 516 AA.
AC Q4LR64;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Phospholipase D/Transphosphatidylase precursor.
GN ORFNames=Bcen2424DRAFT_3286;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
HI2424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
HI2424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AAHL0100025; EAM18523.1; -; Genomic_DNA.
DR SIGNAL.
FT SIGNAL
SQ SEQUENCE 516 AA; 57599 MW; 56B9D87A8D18A60E CRC64;

Query Match 32.4%; Score 881.5; DB 2; Length 516;
Best Local Similarity 40.5%; Pred. No. 1.1e-53;
Matches 214; Conservative 77; Mismatches 203; Indels 35; Gaps 11;

Qy 18 MKTSLSLCLLCSCSSWLPPEERTESHFNTSKPVLNDILQIRHPTHNNGLSDIY 77
Db 2 IRLSWGAISLLMLAACASLPQADRAPTHAFDTDDTRLGVAFRQQAHH-PQODAFH 60
Qy 78 LLDPHPEALAAALIESAHSLSLDLQYIWRNDISGRLLFNLMYLAERGVVRVRLLDN 137
Db 61 LLTDPVDALDARVLLADRAERSLDLQYIWRHDDLTHGELADAVMRAADRGVVRVRLLDL 120
```



```
Qy 140 RGLDLLLALDSHNIENLEPNFVLKRWALGYLTDPPRLNRHMHKSFADNRATILG 199
Db 122 RD-DQVWALLNDHNIENLEPNFYNFRFGLSFLTDFSLNRHMHKSMTADGVAIIVG 180
Qy 200 GRNIGDEYFKVGEDTVFADLDILATGSSVVGSHDFDRIYASHAHNATRIIRSGNIGKG 259
Db 181 GRNIGNEVFSTQVEFGDFVLFGPAVVTATQFDLYWNSDQTQSVASLVSDVETQ 240
Qy 260 LOALGYNDSTSRHALLRYRE-TVQSPPLYQKIQTGRIDMQSVQTRLISDTPAKGLDRDRR 318
Db 241 YALEDLVDVNALEAFPRDGEYDISOGLFEHLKHGTLKLYGAEAVLWYDLDPKVEYRDSQ 300
Qy 319 KPPTAGRLQDALKQPEKSVILVSPYFPTKSGTDALAKLVODGIDVTVTNSLOATDVAA 378
Db 301 ---VAEQALERNVNTSILISPYFVPTAGTIGLKLAEBSGVKVTIVTNSLASNDVFA 357
Qy 379 VHSQVYKRPKLLKAGIKVLELQPNHAPVATKDKGLTGSSVTSIHKATFIIVDGKRIFGS 438
Db 358 VHGYAKYRELLSGITLWIKASAEI--KRWSLIGSQBSLHAKVIVIDDADFVGS 415
Qy 439 FNLDPRSARLNTMGVWIESPKIAEQMERTIADTTPEYAYRVTLDKHNLQWHPATRK 498
Db 416 MNWDPERSAYINTEMAVHITHEPYVKKATVQLNSALPKDAYVVEIREGN-LHWTDLGNKV 474
Qy 499 YPNPEAKLWKRIAKILSLPIEGLL 525
Db 475 YDSEPFESSIWRSIGAWLSGVLPFIEGML 501
```

RESULT 21

```
Q8D861.VIBVU
ID Q8D861_VIBVU PRELIMINARY; PRT; 524 AA.
AC Q8D861
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=V13126;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016807; AAO11447.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 524 AA; 59929 MW; 8C88CBC38C8CADA7 CRC64;
```

```
Query Match 32.0%; Score 871.5; DB 2; Length 524;
Best Local Similarity 39.4%; Pred. No. 5.6e-53;
Matches 211; Conservative 89; Mismatches 197; Indels 39; Gaps 14;
Qy 5 PKIQAMPSETISPMKT---RSLISLLCLLLSCSSWLPPLBERTESRHFNTSKPVLL--- 58
Db 7 PEYLNKPNFFGLGRLPTFIHKALTVCSVLLCACTSS---LENHSP---FDKQPSVQLGYQ 59
Qy 59 -DNILQ---IRHTPHN-NGLSIDYLLDDPHEALARAALIESAEHSLDIQYIWRNDISGR 114
Db 60 ADSRLSAYLNHPQDRENLTAFPLDKGHDALLARLALIEAADKTLDIQYIIFRNDETGQ 119
Qy 115 LLFNLMLYLAERGVRVRLDDNNTRTGLDLLLALDSHPNIEVRLFNPFVLRKWRALGYL 174
```

```
Db 120 LLTWRLFEAQRGVYRVILLDDDMQKRN-DEGLARLSAHPNIQIRLFPHQVYRTARTLAMA 178
Qy 175 TDFPRLNRHMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVVGSHD 234
Db 179 SDFSRLNRHMHKSLTADSVVAIVGGRNIGNEVFSVNSPVEFGDFDLMLYNSVEQTAEQ 238
Qy 235 FDRYWASHSA-----HNATRIIRSGNIGKGLQALGYNDSTSRHALLRYRETVESQSP 286
Db 239 FDLWNSLHAPVIEWLTDNPIPVTE-----ELQA--WLKETQLEQKFTQGRYDFTQLQL 291
Qy 287 YQKIQTGRIDMQSVQTRLISDTPAKGLDRDRRKPPIAGRLQDALKQPEKSVILVSPYFVP 346
Db 292 YQQTDFKSLVWYWGKGQWYDLDPK---VDTQAPQLADNLASLLRTVKDSLVLISPYFVP 348
Qy 347 TKSGTDLAKLVODGIDVTVTNSLOATDVAAVHSGYVYKRPKLLKAGIKVLELQPNHAP 406
Db 349 TERGTQALVEAAHRGVDTITVTNSLASNDVFAVHGWTAKYRQDLVEAGIQLWETKASARI 408
Qy 407 PATKDKGLTGSSVTSIHKATFIIVDGKRIFGSNLDPRSARLNTMGVWIESPKIAEQME 466
Db 409 DS--KWSFTGSSRSSSLHAKVLLIDHRLLFAGSMNWDPRSALLNTTEMAAVIEHPDYV 466
Qy 467 RTIADTTPEYAYRVTLDKHNLQWHPATRKTYPNEPEAKLWKRIAKILSLPIE 522
Db 467 AKLPMGLTNAVYQVRM-KNGEVAWFDHQSWFDESEATVWRKIGAWFAGILFIE 521
```

RESULT 22

```
Q7MMB3.VIBVY
ID Q7MMB3_VIBVY PRELIMINARY; PRT; 554 AA.
AC Q7MMB3
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VV1159.
GN OrderedLocusNames=VV1159;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC93923.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 554 AA; 63341 MW; 8A7980B8096AB6C8 CRC64;
```

```
Query Match 32.0%; Score 869.5; DB 2; Length 554;
Best Local Similarity 38.2%; Pred. No. 8.4e-53;
Matches 207; Conservative 90; Mismatches 194; Indels 51; Gaps 12;
Qy 5 PKIQAMPSETISPMKT---RSLISLLCLLLSCSSW--LPPLERTE-----SR 48
Db 37 PKYLNKPNFFGLGRLPTFIHKALTVCSVLLCACTSFENHSPFDKQPSVQLGYQADSRLSA 96
Qy 49 HFNTSKPVLLDNILQIRHTPHNGLSDIYLLDDPHEALARAALIESAEHSLDIQYIWR 108
Db 97 YLN-----HHQDRENLTAFPLDKGHDALLARLALIEAADKTLDIQYIIFR 143
Qy 109 NDISGRLLFNLMYLAERGVRVRLDDNNTRTGLDLLLALDSHPNIEVRLFNPFVLRKW 168
```

Db 144 NDETGQLTLWLFEEAARGVRVIRLLDDMQKN-DEGLARLSAHPNIQIRLFENHQYRTA 202
Qy 169 RALGYLTDFPRLNRRMHNKSFADNRATILGRNIGDEYFKVGEDTVFADLDILATGSV 228
Db 203 RLTAWASDFSLNRRMHNKSLTADSVVAIVGRNIGNEYFVNSPVEFGDFDLMDYNSV 262
Qy 229 GEVSHDFRYWASHA-----HNATRIIRSGNIGKGLQALGYNDETS-RHALLRYRET 280
Db 263 EQTAEQFDLYNLSHAVPIEWLTDPVTE-----BELQA-WLKETOLEKQKFTQGRYD 315
Qy 281 VEOSPLYXKIQTGRDWSQVOTRLISDPKALQDRDRKPIAGRLQDALQKPEKSVLV 340
Db 316 FTQLQLYQFTKSLVFWGQVWYDLPDK---VDTQEQQLADNLTSLLTFTVDSVLVI 372
Qy 341 SPYFVPTKSGDALAKLVQDGIDVTNLNSIQATDVAHVSGYKVKRPLKAGIKLYEL 400
Db 373 SPYFVPTERGQALVEAARGVEITVITNSLASNDVFAVHWYAKYKRODLVEAGIQWET 432
Qy 401 OPNAVPAKDKGLTGSVTSLSHAKTFIVDGKRFIFGSFNLDPDRSARLNTMGVVIKSPK 460
Db 433 KASARIDS--KWSFTGSRSSLSHAKVLLIDHRLLFAGSMWDPDRSALLNTMAAAVIEHPD 490
Qy 461 IAEQWERTLATTEYAYRVTLDKHRLQWHDPAKTRKTYNPEPAKLVKRIAAKILSLP 520
Db 491 YVOSSEAKLPMLNETNAYQVRM-KNGEVANFQHSQVWFDSEPEATVWRKIGANFAGILP 549
Qy 521 IE 522
Db 550 IE 551

RESULT 23
Q8D800 VIBVU
ID Q8D800 VIBVU PRELIMINARY; PRT; 501 AA.
AC Q8D800;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=VV13197;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AO016807; AO011511.1; -; Genomic_DNA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 501 AA; 56571 MW; BF094CE115F8D30B CRC64;

Query Match 31.7%; Score 862.5; DB 2; Length 501;
Best Local Similarity 38.0%; Pred. No. 2.3e-52;
Matches 197; Conservative 91; Mismatches 198; Indels 33; Gaps 8;
Qy 21 RSLISLLCLLSCSSWLPLEERTESRHNFTSK-PVLLDNLQIRHTPHNGLSDIYLL 79
Db 2 RVLVAVSCFGLACSNPLVTEKKLSYSGYQEGSTLAKYEQLRDETVDKDLNLTGPHSL 61
Qy 80 DDPHEALAAARALTESAHSIDLYIWRNDISGRLLFNLMYLAARGVVRLLDDNNT 139
Db 62 NNGDDALLALIALITSQAHSIDLQYLYFGDDETSKIITWRLYEAAQGVKVRLLDDMQR 121

Qy 140 RGLDLLLLSHDNIEVRLFPNPFVLRKWRALGYLTDFPRLNRRMHNKSFADNRATILG 199
Db 122 RD-DQWALLNDHNIEIRLFPNPFVLRKWRALGYLTDFPRLNRRMHNKSFADNRATILG 180
Qy 200 GRNTGDYFKVGEDTVFADLDILATGSVGVESHDFRYWASHAHNATRIIRSGNIGK 259
Db 181 GRNIGNEVFSFTQVEFGDFDLVLLFGPAVVEATQFDLYWNSDQTPVSAV- 232
Qy 260 LQALGYNDESRHAL-----LRYRE---TVEQSPLYKQIQTGRIDWQSVQTRLIS 306
Db 233 ----SLSDYQYALEDLVDNALVAFPRDGEYDISQLELFEHLKHTGLKLYWEANLWY 288
Qy 307 DTPAKGLDRDRRKPIAGRLQDALQKPEKSVLVSPYFVPTKSGTDLAKLVQDGIDVT 366
Db 289 DLPKVKETRDSQ---VAEQLAERLNNVTYDYLIIISPYFVPTKSGTDLAKLVQDGIDVT 345
Qy 367 LTNSLQATDVAHVSGYKVKRPLKAGIKLYELQPNHAVPATKDKGLTGSVTSLSHAKT 426
Db 346 VTNSLASNDVFAVHWYAKYKRELLTSGITLWEIKASAEI--KQWSLIGSQASLHAKV 403
Qy 427 FIVDGKRFIFGSFNLDPDRSARLNTMGVVIKSPKIAEQWERTLATTEYAYRVTLDKH 486
Db 404 IVIDDADFVGSNMWDPDRSAYINTEMAVHITHPEYKVEATVQLNSALPKDAYVVEIREGN 463
Qy 487 RLQWHDPAKTRKTYNPEPAKLVKRIAAKILSLPIEGLL 525
Db 464 -LHWTDLGNKGVDSPEPSSIRWSIGAWLSGVLPIDEGML 501

RESULT 24
Q8PIM4 XANAC
ID Q8PIM4 XANAC PRELIMINARY; PRT; 520 AA.
AC Q8PIM4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.P., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.W., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL: AO011929; AAM37716.1; -; Genomic DNA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 57436 MW; 61A2DE5E0349A2FA CRC64;

Query Match 31.3%; Score 850.5; DB 2; Length 520;
Best Local Similarity 37.9%; Pred. No. 1.7e-51;
Matches 206; Conservative 80; Mismatches 182; Indels 75; Gaps 15;

QY 24 ISLLCLLCSC-----SSWLPPLLEERTESRHFNFTSKPVLDDNI---LQ 63
DB 12 IALLVLVLASLSLYGYGRFADRGQPVSHVLP-----TAVATIDKVAAPLQ 60
QY 64 IRHTPHNGLSDIYLLDDPHEALARAALIESAEHSLDQYIWRNDISGRLLFNMLYLA 123
DB 61 QAHADQ----TCWVILPNVDAFAVALTARAAGHSLDQYIWHADFTGNLLHNELR 116
QY 124 AERGVRVRLDDNNTRGLDLLLLALDSHPNIEVRLFNPFVLRK---WRALGYLTDFPRL 180
DB 117 ADRGVRVRLDDNNHIGSDSVLAALDSHPLEIRLFPNTRAREGTLMRGVEMVLRMFSI 176
QY 101 NRMHNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFRYWA 240
DB 177 NRMHNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFRYWA 240
QY 241 SHSAHNAIRIIRS-----GNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKI 291
DB 237 SPNALPLSALVTAKPQALEALRGSLDAGWASVRAHP-----YVERLKESPSVRLM 287
QY 292 TG--RIDQSVQVRLISDTPAKG----LDRRRKPPIAGRLQDALKQPEKSVYLVSPYFV 345
DB 288 QGDRPVHVS-QARIVSDPPEKAEGAPQADWMTPLVIGEMAHA----QRELKVISPYFV 342
QY 346 PKSGTDLAKLVODGIDVTVLTNSLQATDVAHVSGYVVKRPLKAGIKLYELOPNHA 405
DB 343 PGDEGMRVIGHLRQNRVRSILTNLSLAANDVVVHSGYAGYRVPLQLQGVRLHEUKP--- 399
QY 406 VPATKDKGLTGSSVTSLSHAKTIFVQKRFIFGSLNLDPRSRALNTMGVWIESPKIARQM 465
DB 400 -MGKPDGSLFGSSGASLHTKAFVDDSSGFGSLNLDPRSRALNTMGVWIESPKIARQM 458
QY 466 ERTL-ADTTPEYAVYVTLQKHLNRLQWHDPAIR--KTYNEPEAKLWKRIAKILSLPIE 522
DB 459 ERLYHNVKVSAPYSYRVTLDR--GELRWHDHDAQPPPEWSREPAASVWRGAAAVMGWLPLE 517
QY 523 GLL 525
DB 518 SOL 520

RESULT 25

Q6MQL7_BDEBA PRELIMINARY; PRT; 530 AA.
ID Q6MQL7_BDEBA PRELIMINARY; PRT; 530 AA.
AC Q6MQL7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative phospholipase D precursor.
GN OrderedLocusNames=Bd0448;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCBI 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842647; CAE78430.1; -; Genomic DNA.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Signal.
FT SIGNAL 1 20 Potential.
SQ SEQUENCE 530 AA; 59501 MW; 4C810BPAFD934F6F CRC64;

Query Match 31.2%; Score 849.5; DB 2; Length 530;
Best Local Similarity 39.3%; Pred. No. 2.1e-51;
Matches 203; Conservative 84; Mismatches 205; Indels 25; Gaps 12;

QY 24 ISLLCLLCSCSWLPPLLEER--TESRHFNFTSKPVLDDNIQIRHTHNNGLSDIYLLDD 81
DB 24 IPILLGLSAGCAS-LPANTNRPYSEALPPDKPTLAQATETQIKGHP---QGSFYPLAS 79
QY 82 PHEALARAALIESAEHSLDQYIWRNDISGRLLFNMLYLAERGVRVRLDDNNTRG 141
DB 80 GLEALVARMAAVVSDRSIDQYIWEIDLGRMLLHEVLRAADRGVRVRLDDLNQSR 139
QY 142 LDDLLALLDSHPNIEVRLFNPFVLRKRALGYLTDFPRLNRRMKNKSTADNRATILGGR 201
DB 140 YEKGLAILDYHPNVEVRMANPFAGRTWKILDAMR-FSTVNRMRMKNKVFVDNQTAVGGR 198
QY 202 NIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFRYWASHSAHNATRIIRSGNIGKGLQ 261
DB 199 NIGDEYFGASEMNFQDFDLWALGPVQVQLSRHFDFTYNSIAIPISVLVK--GFKPTLE 256
QY 262 ALGYNDETSRAL-----LYRETVEQSPLYQKI--QTGRIDWQSVQTRLISDTPAK-GL 313
DB 257 DLQKLKODAAAITEAEKTOYADALKETPIVKKFTHEPMKLYWGKAD--VVMDDPEKPHQ 314
QY 314 DRDRRKPIAGRLQDALKQPEKSVYLVSPYVPTKSGTDALAKLVQDIDVTVLTNSLOA 373
DB 315 DSKDQVNLARQLYPLIEKTEKELILVSPYVPGKGVKFFKHLNDRGVQSLVLTNSLAS 374
QY 374 TDVAHVSGYVVKRPLKAGIKLYELOPNHAPVATKDKGLTGSSVTS--LHAKTFIVDG 431
DB 375 SDVATVSGYGYGRKDLDDQGVSLYELKPN-SPITTPKKNRVGSSFSAGLHGKIFVDR 433
QY 432 KRIFIGSNLDPRSRALNTMGVWIESPKIAEQMERTLADTTPEYAVYVTLQKHLNRLQW- 490
DB 434 KKVFGVSNLDPRSATLNSGEMVVVDSPELAEIMSTNLIAHLRRDSYQVLLNKNLWIK 493
QY 491 --HDPATRKTYNPEPEAKLWKRIAKILSLPIEGLL 525
DB 494 TTDDCGLEHVSFKDPTSSWKRFRKASLSAIFIPESWL 530

RESULT 26

Q4FTL4_9GAMM PRELIMINARY; PRT; 570 AA.
ID Q4FTL4_9GAMM PRELIMINARY; PRT; 570 AA.
AC Q4FTL4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Probable phospholipase D.
GN ORFNames=Psyc 0791;
OS Psychrobacter arcticum 273-4.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Psychrobacter.
OX NCBI_TaxID=259536;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=273-4;
RG NASA Astrobiology Institute;
RA Ayala-del-Rio H.L., Chain P., Ponder M.A., Di Bartolo G., Ivanova N.,
RA Bergholz P.W., Hauser L., Land M., Bakermans C., Rodrigues D.,
RA Klappenbach J.A., Zarka D., Larimer F., Richardson P., Thomas M.F.,
RA Tiedje J.M.;
RT "Complete sequence of Psychrobacter arcticum 273-4.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=273-4;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hamon N., Israni S., Chain P., Di Bartolo G., Ivanova N., Hauser L.,
RA Land M., Larimer F., Pitluck S., Richardson P.,
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000082; AAZ18644.1; -; Genomic DNA.
SQ SEQUENCE 570 AA; 63466 MW; 08CC7DAA9E3CA3BC1 CRC64;

Query Match 31.2%; Score 848.5; DB 2; Length 570;
Best Local Similarity 40.4%; Pred. No. 2.7e-51;
Matches 181; Conservative 87; Mismatches 173; Indels 11; Gaps 5;

Qy 85 ALAARAALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVRVRLLLDD-NNTGLD 143
Db 119 AFASRSILTGATRNIDAQYIWHDDQAGQLMLDLMDAERGVIIVRLLLDDFNNAKFD 178

Qy 144 LLIALDLSHPNIEVRLNPFVLRKWRALGYLTDFPRLNRRMHKNSFTADNRATILGGRNI 203
Db 179 QHLURFASHPNVARIINPLIRKFKSLNFUTGLPRINRRMHKNSMIFDKQITIIIGRNI 238

Qy 204 GDEYFKVGEDTVFADLDILATGVSVVGEVSHDFDRYASHSAHNATRII-----RSGNIGK 258
Db 239 GDGYLSNDKNSQFADLVLLIGKVADIDNSFASVNSAPISFDIQTATLDKGETTFVK 298

Qy 259 GLQALGYNDTSRHALR-VRETVEQSPLVYKIQTGRIQWQSVQTRILISDTPPAKGLDRDR 317
Db 299 GLDKLKTDEKSSSGSLSIYKAAIEDSSIDTDLINKRVPFRWTDQFLSDVGVKLTPTVP 358

Qy 318 RKPIAGRLQDALQKPEKSVLVSPYFVPTKSGTDALAKLVQDIDVTVLNLSQATDVA 377
Db 359 ADTNLHVQLRTLLGSPSKKLTISSYFVPTKGVNTLNALAESGIEIKILTNISFDATDVT 418

Qy 378 AVHSYGKVRKPLKAGIKLYELQPNHAVPATKDKGLTG--SSVTSLHAKTFFVDGKRIF 435
Db 419 AVHSYGQWPSLLRAGVKIYELKSTASEKREKLNKWSQSSTSLHAKAFVDDVQVF 478

Qy 436 IGSFNLDPRLNTEMGVIESPKIAEQMERTIADTTPPEYAVRVTLKRNLRQW--DP 493
Db 479 IGSYVNDPRSANINTGVINDELDLQHDALSDLLQAQYEVKLENGNLQWHTMEK 538

Qy 494 ATRKTYNEPEAKLWKRIAKILSLPIEGLL 525
Db 539 DKVYIDSEPRVAVSHVLTMSWLPIDWLL 570

RESULT 27
Q5QY8_IDILO
ID Q5QY8_IDILO PRELIMINARY; PRT; 464 AA.
AC Q5QY8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=Il1444;
OS Idiomarina loihiensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Idiomarinaceae; Idiomarina.
OX NCBI_TaxID=135577;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=L2-TR / DSM 15497 / ATCC BAA-735;
RX PubMed=15596722; DOI=10.1073/pnas.0407638102;
RA Hou S., Saw J.H., Lee K.S., Freitas T.A., Belisile C., Kawarabayasi Y.,
RA Donachie S.P., Pikina A., Galperin M.Y., Koonin E.V., Makarova K.S.,
RA Omelchenko M.V., Sorokin A., Wolf Y.I., Li Q.X., Keum Y.S.,
RA Campbell S., Denery J., Aizawa S.-I., Shibata S., Malahoff A.,
RA Alam M.;
RT "Genome sequence of the deep-sea gamma-proteobacterium Idiomarina
loihiensis reveals amino acid fermentation as a source of carbon and
energy.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:18036-18041(2004).
DR EMBL; AE017340; AA82284.1; -; Genomic DNA.
GO; GO:0003824; F:catalytic activity; IEA.

DR GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 464 AA; 53557 MW; 59A8BFA13DEB45F7 CRC64;

Query Match 31.1%; Score 846; DB 2; Length 464;
Best Local Similarity 40.2%; Pred. No. 3e-51;
Matches 186; Conservative 76; Mismatches 173; Indels 28; Gaps 6;

Qy 78 LLDDPHEALAARAALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 137
Db 15 LLENAHRALIERVRLIRQARHIALQYILWRPSSGLTLLKELLDAVERGVRVDDLLDDL 74

Qy 138 NTRGLDLLLLALDLSHPNIEVRLNPFVLRKWRALGYLTDFPRLNRRMHKNSFTADNRATI 197
Db 75 HSKPLEPLDLSSRANFNIRLNFPPKREWRNLNLTDFKRMNRMHKNSMLVDAEYGI 134

Qy 198 LGGRNIGDEYFKVGEDTVFADLDILATGVSVVGEVSHDFDRYASHSAHNATRIIRSGNIG 257
Db 135 VGGRNVGNEVFGTHAGQLFSDLDVIAKGKIVKTLDDWQYW-----NCTLSVAVNAIS 188

Qy 258 KGLQALGYNDTSRHALRYRETVEQSPLVYKIQI-----QTGRIDWQSVQTRILISD 307
Db 189 SSKPLRFLND-----TWGQYKQPELRDLHQFVLSDEHLTDTPQSDAETWYQATATVVD 243

Qy 308 TPAKGLDRDRRKPIAGRLQDALQKPEKSVLVSPYFVPTKSGTDALAKLVQDIDVTVL 367
Db 244 DPLKALPDSKPSRLVTQTLTGLAKQSIVLVSPYFVPTGTGVKELENLSEQGVKVRVL 303

Qy 368 TNSLQATDVAHVSHGVYKVRKPLKAGIKLYELQPNHAVPATKD-KGLTGSVTSLSHAKT 426
Db 304 TNSLAVTDVPAVAGVQRHRRLLLSGVLFDEKRTTEKPLRQDRVPYFRSSASSLHAKT 363

Qy 427 FIVDGKRIFIGSFNLDPRLNTEMGVIESPKIAEQMERTIADTTPPEYAVRVTLKRN 486
Db 364 VTVDSEKVFVGSFNFDPRSAQINTESGLLIESAEARINAMFEHEHPLRSYEALRNFY 423

Qy 487 RLQWHD-----PATRKYTPNEPEAKLWKRIAKILSLPIEGLL 525
Db 424 KLYWLDKSKIPALR--LYKEPGAGIWRRLAVMLTARLPVDHLL 464

RESULT 28
Q4UWU7_XANCP
ID Q4UWU7_XANCP PRELIMINARY; PRT; 520 AA.
AC Q4UWU7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cardiolipin synthase.
GN ORFNames=XC.1408;
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-F., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP00050; AA48476.1; -; Genomic DNA.
SQ SEQUENCE 520 AA; 57165 MW; 8849CFED240410A3 CRC64;

Query Match 31.0%; Score 843; DB 2; Length 520;
Best Local Similarity 40.4%; Pred. No. 5.8e-51;
Matches 196; Conservative 71; Mismatches 174; Indels 44; Gaps 12;

```
QY 62 LQIRHTPHNGLSDIYLLDDPHEALAAARAAALIESAEHSLDQYYIWRNDISGRLLFNLMY 121
Db 59 LQQAHA-----GKTGMVILSDNIDAFVRAALTARAAGSLDQYYIWHADFTGNLLHNELL 114
QY 122 LAAERGVVRLLDDNNTRGLDDLLALDHPNIEVLPNPFVLRK---WRALGYLTDFF 178
Db 115 RAADRGVVRLLDDNNHGHSDVLAALDHPLEIRLFPNTRAREGTLMRGVEMVLRMF 174
QY 179 RLNRHMKNKFTADNRATILGGRNIGDEYFKVGDVTFADLDILATGTVVGEVSHDFDY 238
Db 175 SINRRMKNKAWIADGRIAVVGGNVGDEYFDAARDTNFMDMDAALIGPTVQSEQVFDAY 234
QY 239 WASHSAHNATRIIRS-----GNICKGLQALGYNDETSRHALLRYRETVEQSPLYQK 289
Db 235 WNSPNALPLAALVTAKPOALEQLRGLSDAGMASVGAHP-----YVERLRQSPSVHA 285
QY 290 IQTGR--IDWQSVQTRILSDTPAKG---LDRDRKPPIAGRLQDALQKPSVYLVSPY 343
Db 286 LMQGNRPVHWVQ-QARIVADPPEKAEGAPPDWMTPILIGEMANA-----RRELKLIISPY 340
QY 344 FVPTKSGTDALAKLVQDGDIVTLTNSLQATDVAHVHSGYVKYRKPLLKAGIKLYELQPN 403
Db 341 FVPGDDGMRWIGGLRKDRVSVLTNSLAANDVVAHVHSGYADYRIPLLQQGVRLHELKP- 399
QY 404 HAVPATKDKGLTSSVTSLSHAKTFIVDGKRIPIFGSNLDPRSARLNTMGVVISPKIAE 463
Db 400 ---MGRPDGSLFGSSGASLHTKAFVVDSDSAGIFGSLNDPRAMNLTMGLLFEDRTVTA 456
QY 464 QMERTL-ADTTPYAYRVTLDKHNLQWHDPATR--KTPYNEPEAKLWKRTAAKILSLP 520
Db 457 ELERLYNHKVSAPVSYRVTLQ-QGLRWHDDAAQPPAVWTREPAASVWRRGAAGVGLWLP 515
QY 521 IEGLL 525
Db 516 LDSQL 520

RESULT 29
Q8P7A5_XANCP
ID Q8P7A5_XANCP PRELIMINARY; PRT; 520 AA.
AC Q8P7A5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cis;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorzy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Subetal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
```

```
DR EMBL; AE012383; AAM41978.1.; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
SQ SEQUENCE 520 AA; 57165 MW; 8849CFED240410A3 CRC64;

Query Match 31.0%; Score 843; DB 2; Length 520;
Best Local Similarity 40.4%; Pred. No. 5.8e-51;
Matches 196; Conservative 71; Mismatches 174; Indels 44; Gaps 12;

QY 62 LQIRHTPHNGLSDIYLLDDPHEALAAARAAALIESAEHSLDQYYIWRNDISGRLLFNLMY 121
Db 59 LQQAHA-----GKTGMVILSDNIDAFVRAALTARAAGSLDQYYIWHADFTGNLLHNELL 114
QY 122 LAAERGVVRLLDDNNTRGLDDLLALDHPNIEVLPNPFVLRK---WRALGYLTDFF 178
Db 115 RAADRGVVRLLDDNNHGHSDVLAALDHPLEIRLFPNTRAREGTLMRGVEMVLRMF 174
QY 179 RLNRHMKNKFTADNRATILGGRNIGDEYFKVGDVTFADLDILATGTVVGEVSHDFDY 238
Db 175 SINRRMKNKAWIADGRIAVVGGNVGDEYFDAARDTNFMDMDAALIGPTVQSEQVFDAY 234
QY 239 WASHSAHNATRIIRS-----GNICKGLQALGYNDETSRHALLRYRETVEQSPLYQK 289
Db 235 WNSPNALPLAALVTAKPOALEQLRGLSDAGMASVGAHP-----YVERLRQSPSVHA 285
QY 290 IQTGR--IDWQSVQTRILSDTPAKG---LDRDRKPPIAGRLQDALQKPSVYLVSPY 343
Db 286 LMQGNRPVHWVQ-QARIVADPPEKAEGAPPDWMTPILIGEMANA-----RRELKLIISPY 340
QY 344 FVPTKSGTDALAKLVQDGDIVTLTNSLQATDVAHVHSGYVKYRKPLLKAGIKLYELQPN 403
Db 341 FVPGDDGMRWIGGLRKDRVSVLTNSLAANDVVAHVHSGYADYRIPLLQQGVRLHELKP- 399
QY 404 HAVPATKDKGLTSSVTSLSHAKTFIVDGKRIPIFGSNLDPRSARLNTMGVVISPKIAE 463
Db 400 ---MGRPDGSLFGSSGASLHTKAFVVDSDSAGIFGSLNDPRAMNLTMGLLFEDRTVTA 456
QY 464 QMERTL-ADTTPYAYRVTLDKHNLQWHDPATR--KTPYNEPEAKLWKRTAAKILSLP 520
Db 457 ELERLYNHKVSAPVSYRVTLQ-QGLRWHDDAAQPPAVWTREPAASVWRRGAAGVGLWLP 515
QY 521 IEGLL 525
Db 516 LDSQL 520

RESULT 30
Q6G185_BAROU
ID Q6G185_BAROU PRELIMINARY; PRT; 523 AA.
AC Q6G185;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=B001680;
OS Bartonella quintana (Rochalimaea quintana).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=803;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Toulouse;
RX PubMed=15210978; DOI=10.1073/pnas.0305659101;
RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
RA Canback B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,
RA La Scola B., Holmberg M., Andersson S.G.E.;
RT "The louse-borne human pathogen Bartonella quintana is a genomic
RT derivative of the zoonotic agent Bartonella henselae.";
```

RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacteri
RT *Mesorhizobium loti* ";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003017; BAB54862.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.


```

DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN Cardiolipin synthetase family protein.
OS OrderedLocusNames=mlr8077;
OC Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338 (2000).
DR EMBL; BA000012; BAB53717.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
KW SEQUENCE 466 AA; 52117 MW; 693301E508A0A788 CRC64;
SQ
Query Match 29.6%; Score 804.5; DB 2; Length 466;
Best Local Similarity 41.0%; Pred. No. 2.6e-48;
Matches 193; Conservative 61; Mismatches 180; Indels 37; Gaps 10;
QY 72 GLSDIYLLDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVVR 131
DB 16 GKSGLLISDNYDAFAARVLAARGAARTLDMYLLWHDDHGRLLLOEVVRAAQRGVVR 75
QY 132 LLDDNNTRGLDLLLLALDSDHPNIEVRLFNFPVLNRK---WRALGYLTFDFPRLNRRMHKS 188
DB 76 MLDDVNPXKSDAAYLALSNNHPIELKLFNPSGIRARGLARGAEVLLRLPALTRWHNKA 135
QY 189 FTADNRATILGRRNIGDEYFKVGEDTFVADLILATGSSVGEVSHDPRYWAHSAHNAT 248
DB 136 WIADDNTAIYVGGVNGVDYDAAE-TNFRDLMLLLGPAVQQAQIFQTFEWCQDA---- 190
QY 249 RIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYKIQ-----TCRIDW 297
DB 191 -----KPIAELGAAAGSGRAPYFEGREEKTESTLLSGIRDKGSIABFISASSNVHW 241
QY 298 QSVQTRLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFPVPTKSGTDALAKL 357
DB 242 VE-RVRVISDPPEK-VRGWRPRSLMKELLPPIQSARKVEIVSPYFIPGKKGSKILGDL 299
QY 358 VQDGIDVTVLNSLQATDVAHVSGYVKYKPKLLKAGIKLYELQPNHAVPATKDKGLTGS 417
DB 300 VDDGVQAVLTNSLAATDVAHVGAANYRKRLLRMGVQLFELQPFPSRQPKI---SVFGS 356
QY 418 SVTSLHAKTFIVDGKRIFGSNLDPSARLNTMGVVIESP-K-IAEQMERTLADTTPEY 476
DB 357 KGASLHAKTSVONRIGFVGSFNPDPKSVLSNEMGVLFEDENLVAELRHRFKEIAPEA 416
QY 477 AYRVTLDDKHNRLQWH--DPATRTYTPNEPEAKLWKRIAAILSLPIEGLL 525
DB 417 SYRLEL-KNEVLRWHGSDGRLQTYTTEPEAAWFRRLAALVRLPIESQL 466
RESULT 35
Q6FB53 ACIAD PRELIMINARY; PRT; 520 AA.
AC Q6FB53;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Putative phospholipase D protein.
GN OrderedLocusNames=ACIAD1881;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Orsion L.N., Weissbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779 (2004).
DR EMBL; CR543861; CAG68710.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
KW SEQUENCE 520 AA; 60385 MW; 1DD312966EA749A0 CRC64;
SQ
Query Match 29.2%; Score 794; DB 2; Length 520;
Best Local Similarity 37.7%; Pred. No. 1.7e-47;
Matches 186; Conservative 76; Mismatches 177; Indels 54; Gaps 11;
QY 66 HTPHN--NGLSDIYLLDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLA 123
DB 49 HTDENLQKGLTAVIPLDDAFISATRIYLIRHAKHQIDLOYYIWHNDFVQQLILNELKA 108
QY 124 AEGVRVRLLLDNNTRGLDLLLLALDSDHPNIEVRLFNFPVLNRKARGALGYLTFPRLNRR 183
DB 109 ADRGVKIRLLIDQNGTRLDSTFQALLQHPLEIKVFENPYKFRKRVDFLFRPNQINHR 168
QY 184 MNKSTADNRATILGRRNIGDEYFKVGEDTFVADLILATGSSVGEVSHDPRYWAHSHS 243
DB 169 MNKLLIADNTIATVGTGRNISSEYFDASVQFQTFDLDILFYTAASQAEKVQTFWDSPL 228
QY 244 AHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQ--SPLYKIQITGRID----- 296
DB 229 SVTTQKLITKGPNDLLQLRS-----KYKE-IERVDTPTEKINEAQLENEAL 276
QY 297 -----WQSVQTRLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFPVPTKS 349
DB 277 KQRYSVKW--AKAHFVADSPKKITHGHAIGNELIYNQMFSGIMHPQKHLELVSAFYVPTQA 334
QY 350 GTDALAKLVQDGDIVTLNSLQATDVAHVSGYVKYKPKLLKAGIKLYELQ----- 402
DB 335 GTQYLSNLQQDVVKIRVLTNSFAANDVAIVHATYSQTRKELIKNGIQLYEFKPIIDRKEP 394
QY 403 -----NHAVPATKDKGLTGSVTSLSHAKTFIVDGKRIFGSNLDPSARLNTMGV 455
DB 395 TWYKMTGRVPA-----KGKSSSLHAKFFDVGCK-VFIGSFNDFPSAYLNTVEGLV 447
QY 456 IESPKIAEQMERTLADTTPEYAYRVTLDDKHNRLQWH-DPATRTYTPNEPEAKLWKRIA 512
DB 448 VESEDLQNETTKVLDEYLPYATYQLKLDKNGNLVWLDHNDGQTIEYHHPDPTTRFORM 507
QY 513 AKTSLSLPIEGLL 525
DB 508 MKAVSYLPYIEWMM 520
RESULT 36
Q9HTP4_PSEAE PRELIMINARY; PRT; 529 AA.
ID Q9HTP4_PSEAE PRELIMINARY;
AC Q9HTP4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
```



```
QY 423 HAKTFVUDGKRIFTSGNLDPRSARLNTMCMGVIESPKIAEQMERTIADTT-PEYAYRVT 481
DB 431 HAKAYVIDRKILVIGSMNLDPRSARLNTMCMGVIESPKIAEQMERTIADTT-PEYAYRVT 490
QY 482 LDKHN-----RLQWHDPAIR--KTYPNPEAKLWKRIAAKILSLPIEGLL 525
DB 491 LAPDTPGAQTTGAGAPAWPLVWTEIADGQVRYTSDPNAGFYRNLLTGLCLLPIDDQL 550

RESULT 38
Q88CA5_PSEPK
ID Q88CA5_PSEPK PRELIMINARY; PRT; 517 AA.
AC Q88CA5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=PP5276;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae;
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohelsel J., Straetz M., Helm S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RA "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016793; AAM70841.1; -; Genomic_DNA.
DR TIGR; PP5276; -;
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
SQ SEQUENCE 517 AA; 57574 MW; 107E07900AA35751 CRC64;

Query Match 28.4%; Score 773; DB 2; Length 517;
Best Local Similarity 37.8%; Pred. No. 5.2e-46;
Matches 199; Conservative 83; Mismatches 210; Indels 34; Gaps 15;

QY 21 RSLISLCLLL--CSCSWLPPLEERTESRHFTSKPVLDDNLIQIRHTPHNGLSDIYL 78
DB 5 RAUPLLLVLLVAGCTSIAP---RETSQALPAHESAFGRSVL--RQAAPYGRSGFRL 59
QY 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVVRVRLDDNN 138
DB 60 LPNSNEAFRAAEILIRNAQASIDLIQYIVHDGLSTRALVHELLRAADRGVVRVRLDDTT 119
QY 139 TRGLDDLLLDLSDHNPVLEVPFLRK---WRALGYLTDPPRLNRMHKNSFTADNR 194
DB 120 SDGLDVTVMGLDAHPNTHIRVFNPLHLGRSTGVTRAVGRFLNLSRQHRHMKLFLVDNS 179
QY 195 ATILGGRNIGDEYFKVGEDTVFADLDILATGVSUVGVSHDFDRYWASHSAHNATIRSG 254
DB 180 MAIVGGNIGDEYFADPNLNFDTIDLLGVPVAEQLGHSFDQVNSALSRPITDFLWHD 239
QY 255 NICKGLQALGVNDETSHALLRVRETVEQSPLYQKIQT----GRID-WQS----VQTRLI 305
DB 240 PDANDLRASQRLEVS---LAKAR--TQRKALYDRLMAYQSPRLDWRNELIWAHAQAL 294
```

```
QY 306 SDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYSPYFVPTKSGTALAKLVQDGD 363
DB 295 WDAPSKVLADEPDQPLLSSQQLAPDLANVHRELVLASAYFVGPGLLVTGRADAGVS 354
QY 364 VTVLNLSQATDVAAVHSGYKVRKPLLKAGIKLYELOPNHAPVATKDKGLTGSSVTSLSH 423
DB 355 VKLLTNSLEATDVPVHGGYAPYRRALLEHGVQLYELRRQPGDPSAGRLSFRGSSDSLH 414
QY 424 AKTFIVDGKRIFTSGNLDPRSARLNTMCMGVIESPKIAEQMERTIAD--TTPEYAYRVT 481
DB 415 TKAIVFDRRKTFISGFNDFPRSVLWNTVEGVLDVDSPELAE-YTRELAQQGNAPALSQVVK 473
QY 482 LDKHNRLQW--HDPATRKTYPNPEAKLWKRIAAKILSLPIEGLL 525
DB 474 L-IGNKLWATDNGQRHVLTSFPGGIWRRFNVAWISKAVGLEKML 517

RESULT 39
Q4LJD9_9BURK
ID Q4LJD9_9BURK PRELIMINARY; PRT; 540 AA.
AC Q4LJD9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Phospholipase D/Transphosphatidylase precursor.
GN ORFNames=Bcen2424DRAFT_1235;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
RX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHL01000070; EAM16198.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 540 AA; 59105 MW; EA4326D0970D659D CRC64;
```

```
Query Match 28.4%; Score 772.5; DB 2; Length 540;
Best Local Similarity 36.8%; Pred. No. 6e-46;
Matches 198; Conservative 83; Mismatches 208; Indels 49; Gaps 18;

QY 19 KTRLSLSLCLLCSSWLPPLPE-ERTESRHFTSKPVLDDNLIQIRHTPHNGLSDIY 77
DB 21 QARLVVCALLPLAACATHTPATSIDRPVSHALSATATPLRALAAPEAAH-PGQSGFR 79
QY 78 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVVRVRLDDNN 137
DB 80 LLADGATALQMRALARAATKTLDMQYIATEDTTCKLLGALYAADRGVVRVRLVDDL 139
QY 138 NTRGLDDLLLDLSDHNPVLEVPFLRK----WRALGYLTDPPRLNRMHKNSFTADN 193
DB 140 NFHDIDRVMAALNTHQNIIEIRVFNPEFGASQGRMMERTANFFTRIDSFTRMHKMIADN 199
QY 194 RATILGGRNIGDEYFKVGEDTVFADLDILATGVSUVGVSHDFDRYWASHSAHNATIRRS 253
```

